Študentski filozofski zbornik 2018

PHILOSOPHY STUDENTS' COMPENDIUM 2018

Zbornik študentov oddelka za filozofijo Filozofske fakultete Maribor

Maribor, 2019

ŠTUDENTSKI FILOZOFSKI ZBORNIK 2018

PHILOSOPHY STUDENTS' COMPENDIUM 2018

Zbornik prispevkov z mednarodnega študentskega simpozija/Compendium of works from the international philosophy symposium

Urednik/Editor:

Tadej Todorović, magister filozofije in magister anglistike

Recenzenti/Reviewers:
red. prof. dr. Bojan Borstner
izr. prof. dr. Janez Bregant
izr. prof. dr. Friderik Klampfer
ml. razisk. Alen Lipuš
red. prof. dr. Danilo Šuster

Risba na naslovnici/Cover drawing: Tina Ritlop, dipl. um. zgod. (UN) in dipl. ang. (UN)

Lektor/Proof-reader:

Tadej Todorović, magister filozofije in magister anglistike

Oblikovanje in prelom/Design and layout: Tadej Todorović, magister filozofije in magister anglistike

Izdala/Published by: Oddelek za filozofijo, Filozofska fakulteta, Univerza v Mariboru in Društvo za analitično filozofijo in filozofijo znanosti

Kraj/Published in: Maribor, Slovenija

Leto izida/Year of publication: 2019

ISSN 2464-0468

Zbornik je dostopen na spletnem naslovu/The compendium is accessible at:

https://daf.splet.arnes.si/studentski-zbornik/ .

Table of Contents

Preface

The collection of papers before you was produced as a result of proceedings from the *International Students' Symposium 2018*, which was organized by the Department of Philosophy, University of Maribor, and the *Slovenian Society for Analytic Philosophy* (*DAF*) in May 2018.

The symposium, which celebrated its fourth iteration in 2019, has been growing steadily since its inception in 2016. We are honoured that there has been such an amazing response to our invitation from colleagues from other universities, and we hope that it will continue to grow in the future with even more excellent presentations and discussions, and that it will continue to connect young academics from various universities and help in spreading analytic philosophy throughout the globe.

The collection, which is comprised of five select papers presented at the symposium, covers a range of philosophical areas. Daniele Mario Cassaghi (University of Milan) and Aleksa Čupić (University of Belgrade) deal with topics within philosophy of mind; Niko Šetar's (University of Maribor) article belongs to the field of ethics; and Jelena Pavličić's (University of Belgrade) and Lalit Saraswat's (Indian Institute of Technology Bombay) article to epistemology.

I would also like to use this opportunity to give my thanks to the Department of Philosophy and the *Slovenian Society for Analytic Philosophy* for their help in organising the symposium and in publishing this compendium, to all those who volunteered to review the papers, to all those who helped in organising the symposium in some other way, and to Tina Ritlop, for kindly allowing us to use her work as our cover.

Hopefully, the articles in front of you will furnish you with an idea of what philosophical issues those at the start of their respective academic careers care about, and, more importantly, how they do philosophy.

Tadej Todorović

Daniele Mario Cassaghi

University of Milan

Phenomenal Intentionality: Original or Naturalisable?

Abstract

According to Hartry Field, the main challenge for every naturalistically respectable picture of the mind is to account for mental properties in naturalistically adequate terms. Thus, we need a naturalisation of experiential properties (the 'phenomenal character', 'What-it-is-like-ness' of experience) and intentional properties; namely, the ability of a wide array of mental states to be *about, of, refer to* something else from themselves.

In this paper, I will give a general overview of the two main approaches trying to understand the relation between experiential and intentional properties. These are the *natural-externalist research programme* and the *phenomenal intentionality theory*. I will assess their strengths and weaknesses, especially in relation with the naturalisation issue by Hartry Field. Whilst the naturalist-externalist approach is in a better position to meet the challenge by Field, it has several pitfalls that are easily overcome by the phenomenal intentionality theory, which is not as able to address Field's issue in turn.

Finally, I will assess the only proposal that tries to unify the two approaches, the *self-representational higher-order theory* by Uriah Kriegel. I will conclude by stating some brief remarks on some difficulties of this conciliatory proposal.

Keywords: representationalism, intentionalism, original intentionality, phenomenal intentionality, naturalistic theories of intentionality, self-representational high-order theory

0   Introduction

In his 1978 seminal work "Mental Representation", Hartry Field described the main challenge for every naturalistically respectable picture of the mind: the requirement of a plausible account of mental properties in naturalistically adequate terms. The two mental properties he referred to are the same philosophers talk about nowadays: experiential properties (the 'phenomenal character', 'What-it-is-like-ness' of experience) and intentional properties, namely the ability of a wide array of mental states to be *about, of, refer to* something else from themselves.

Although Field's milestone work was about intentionality in abstraction from its relation to experience, it is a controversial issue whether it is possible to treat the two issues separately. The problem of intentionality and its relations to the phenomenal character of experience are the two main concerns of this paper. My focus will be on *how it is possible for mental states to exhibit intentionality* and on what is the role of consciousness in answering this question. I shall start by defining the 'problem of intentionality' and show its connection with semantical issues in philosophy of language and with consciousness. In the second section, I shall make an overview of the first family of approaches to the problem of intentionality and consciousness: the *NERP programme*. In the third section, I shall describe the reverse approach to intentionality, the *phenomenal intentionality*, and present the only view that arguably conciliates the two main approaches: the SHOT account by Kriegel (2011). I shall conclude by arguing that the conciliatory view is not promising.

1   Intentionality, Experience and Originality

Beliefs, desires and perceptions are paradigmatically intentional states. I may hold beliefs *referring to* the capital of China, desires *of* soft drinks, and perceptions about lamps. Unfortunately, the notion of intentionality does not come in a unitary fashion. This notion is meant to capture a cluster of arguably related features of mental states:

> 1) their ability to '*mis*represent' (I may believe that the capital of China is Maribor);
>
> 2) their ability *to refer* (the '*aboutnes/of-ness'* shown above);
>
> 3) their *directedness* (my experience of a lamp intuitively involves a relation with the lamp); and

4) their aspectual shape (I may believe that Hesperus is Venus, without believing that Phosphorous is Venus, since 'Hesperus', 'Phosphorous' and 'Venus', are indeed different ways, i.e. *aspects*, in which the same object is represented) (Dretske 1994: 28-34).

All these properties captured by the concept of intentionality strongly suggest that intentional states *really are* representational states.[1] Representations, like pictures, maps or *linguistic expressions,* share all the features 1-4; they can be true, false, accurate or inaccurate. They refer to the object they represent and so they are related to them. They often report just an aspect of the object in their content: I may represent 'the morning star' by pronouncing 'Hesperus', but I may not represent 'the evening star', even though Hesperus is one and the same as both 'the morning star' and 'the evening star' (cf. Frege, 1948).

Following this suggestion, it seems that the properties characterising intentionality, namely the properties providing troubles to a naturalistically adequate theory of mind according to Field (1978), are representational properties; i.e. semantical properties linked to the representational content of intentionality. The challenge is to provide *psychosemantics*. In this respect, the 'problem of Intentionality' sails close to philosophy of language: it becomes how *i)* it is possible for mental states to have a content (i.e. to *refer to* something else), and (if Field (1978) is right) *ii)* how it is possible to account for the meaning of our mental states in naturalistically kosher terms.

We have said that perceptual states enjoy intentionality. In addition to this, they are also thought to be paradigmatic cases of experiential states, that is, states endowed with a phenomenal character. According to a popular view called '*intentionalism*' (Byrne, 2001) the phenomenal properties of an experience are nothing over and above the properties of its content. Intentionalism in every form claims that sameness in phenomenal character must be due to sameness in content.

Accepting intentionalism binds the 'problem of intentionality' to an account of phenomenal properties. If the phenomenal character of an experience is nothing over and above its content, then a good account of the semantical properties of the content should give us a good account of phenomenal properties in turn. On the other hand, a good understanding of our phenomenal properties

---

[1] This is less controversial for beliefs and desires. It is disputable for perceptions, where the debate between intentionalism (Byrne 2001) and naïve realism (Brewer 2011) is still open nowadays. For the purposes of this paper, I will assume representationalism is true. I will say something more about the representational character of perception later on.

may be a starting point for the account of content. The rest of the paper will be about assessing these two alternatives.

Finally, the last idea to take into consideration is the *originality thesis*. The idea behind it is that along with the many analogies between mental states and ordinary representations, there is still an element of strong disanalogy. The shared intuition among philosophers accepting the originality thesis (Searle, 1992; Kriegel 2011; Horgan, 2013) is that the mind has a special place in the domain of intentionality; our mental states enjoy intentionality in an intrinsic, original manner. Endorsers of the originality thesis take for granted that it is possible to *treat* many things, from artefacts like maps to 'natural signs' like the rings in the tree log, *to refer to* something else. The pirate map means that the treasure is buried in a certain place. The rings in the tree log mean that the tree is of a certain age. The crucial point is that all the other entities we attribute intentionality to (such as maps or other signs) count as intentional only because they enjoy a certain relation toward the original form of intentionality (namely, *our* intentionality). In other words, our mental states are intentional in a way that is metaphysically prior to all the other things we may assign contents to.

In the next section, I will assess the first family of attempts to solve the problem of intentionality. They reject the originality thesis.

2   The Natural-Externalist Research Programme

Following Kriegel (2013a), we may call the first attempt to explain how mental states comes to be intentional the *natural-externalist research programme* (NERP). Proponents of NERP take the challenge by Field seriously – the only acceptable theory of intentionality should provide a naturalistic account of the semantics of our mental states. Therefore, their theories aim to *reduce* intentional properties to certain classes of natural properties in the environment. The NERP orthodoxy requires these properties to be present in the explanations provided by natural sciences, especially biology. Moreover, if intentionalism is true, as many proponents of NERP claim (Dretske, 1994; Tye, 1995; in a certain way Lycan, 1996), it is possible to naturalise consciousness 'for free' once the problem of intentionality is solved (Kriegel, forthcoming).

The major divide within NERP concerns the relevant class of natural properties. Dretske (1981) takes intentionality to consist of reliable *informational co-variance* from some inner states S of the organism (e.g. the vehicle of the representation of a cow) and properties in the environment (e.g. the

property of *being a cow*). This co-variance is *counterfactual* supporting: if there were nothing instantiating the property of being a cow in the environment, the probability of S to occur would get lower. In this sense, Dretske claims, S provides *information* about cows.

Millikan (1989) considers mental representations as *proper* functions; namely, contents are acquired because natural selection favoured their fixation. Mental contents were selected in our ancestors because their 'consumption' had some evolutionary advantages. Here is an example: representing a predator *as a predator* (when there really is a predator) may trigger the consumer system to cause the behaviour of fleeing. This behaviour increases fitness and favours the selection of the content *predator. Predator* acquires the function of triggering the behaviour of fleeing.

Fodor (1990) takes intentionality to be nested into causal relations connecting properties in the environment and inner states of the organisms. A mental state represents *cow* if it is caused by the occurrence of the property of *being a cow* in the environment.[2]

Despite the above differences, NERP's general claim is that mental states have the content they have in virtue of one of natural relations (informational co-variance, evolutionary history, causality) they bear toward properties in the environment. NERP's theorists give up the idea of the originality thesis, because, ultimately, these natural relations are to be found in the environment and not in our mind. This makes NERP theories acceptable by any naturalist and hence tailor-made to solve the problems described by Field.

The main challenge for NERP theorists is the *disjunction problem* (Fodor, 1990), according to which naturalistic theories of intentionality cannot secure both determinate and potentially false representations. To understand this, think about the representation of a fly through the eyes of a frog snapping its tongue to get food. Plausibly, little black dots are also sufficient to erroneously trigger the frog's tongue. What content is fixed for the frog's representation? Since in the frog's environment the correlation between little black dots and flies is dramatically high (there are very few dots that are not also flies), both 'flies' and 'little black dots' may be selectively advantageous for the frog. Both

---

[2] The peculiarity of Fodor's theory, which goes outside the scope of this essay, is the asymmetric-relation dependency. Briefly: if there is something else, for example a bear, that may cause the representation of cow, it is so just because the causal chain between the bear and the representation of cow is dependent on the causal chain connecting the cow and the cow-representation. Since the dependency does not also work the other way round, the content of the cow-representation is a cow and not a bear.

contents may informationally co-vary with representations occurring in the frog, and little black dots are causally sufficient to elicit those representations. So either the frog represents something indeterminate in-between 'flies' and 'little black dots,' or the fixed content is a disjunctive one: 'Little black dots V flies'. However, now we have to give up the idea that the representation of a little black dots is *false*, since it corresponds to one of the disjuncts.[3]

The consequences of failing to solve the disjunction problem are dramatically evident once we focus on the relations between experience and intentionality. In optimal conditions, our contents of experience are not undetermined; when I see a fly on my hand, I see it as a fly, not as something indeterminate in-between flies and little black dots, nor do I see the disjunction fly or little black dot. If we hold that experiential properties are determined by intentional properties, then the disjunction problem must be addressed.

3   The Phenomenal Intentionality Theory

The main adversaries of NERP endorse the *phenomenal intentionality theory* (PIT) (Horgan and Tienson, 2002; Loar, 2003). According to PIT, it is the phenomenal character that grounds the semantics of our mental states. The standard argument in this direction is a variant of Putnam's *brain-in-a-vat* experiment. Suppose that the brain-in-a-vat is tied to a machine designed to induce in the brain-in-a-vat exactly the same stream of consciousness, the same phenomenology, that I enjoy. When I see a lemon, my brain-in-a-vat twin also perceives a lemon. When I listen to a song, my brain-in-a-vat twin also enjoys the same auditory representation as I do. My perceptions are *about* lemons and songs, and the same is true for my brain-in-a-vat twin. Since the only common factor between me and my brain-in-a-vat twin is the phenomenology we share, it is plausible to claim that sameness of content is determined by sameness of phenomenology, especially if we understand content as truth-conditions, following the orthodoxy in analytic philosophy. The brain-in-a-vat's states have the same contents as mine, but the truth-conditions cannot be dependent on external factors: there is nothing in the brain-in-a-vat environment. Thus, in this view, nothing but phenomenology can contribute to content determination. The conclusion is a narrow view of content, fully grounded in phenomenology (Farkas, 2008). The argument assumes intentionalism since the same contents are matched to the same phenomenal characters.

---

[3] Fodor (1990) tries to fix this by adding the asymmetric-dependency clause (see Footnote 2). However, this move make the theory circular: you cannot determine the right causal chain without assuming the content you wish to fix. See Adams and Aizawa (1994).

To sum up: PIT theorists address the problem of intentionality in the following way: some mental states, in particular perceptual states, are intentional because it is the phenomenal character of the experience that determines its content, as is in the case of the brain-in-a-vat. PIT accepts the originality thesis at heart: intentionality nested into phenomenal character does not depend on any other factors external to the subject (Horgan, 2013). Finally, it is open for PIT theorists to consider other mental states, like beliefs and desires, to also be phenomenal states (a doctrine known as 'cognitive phenomenology'), or to let them just be *derivatively* intentional, namely intentional in virtue of some relation toward mental states of the former kind. Going back to the problem of intentionality, as formulated by Field, we can say that, *according to PIT, a mental state acquires its content <u>either</u> in virtue of the phenomenal character it enjoys (original intentionality) <u>or</u> in virtue of some relation the mental state bears toward other originally intentional mental states.*

The main advantage of PIT is the fact that the disjunction problem has an easy solution: it makes a phenomenal difference when representing a fly instead of a little black dot. Thus, there is no doubt on whether I am representing a fly or a little black dot. The contents of phenomenal intentionality are *determined* (Horgan and Graham 2013). The *originality thesis plays a key role*: the lack of external factors at determining contents erases the risk of indeterminacy (compare to NERP). Again, remember what is at issue for my brain-in-a-vat twin. This seems to suggest that phenomenal character is able to generate *misrepresentations* in the way required for the solution of the disjunction problem.

The obvious concern with PIT is giving up meeting Field's constraints about a naturalistically adequate theory of intentionality. Indeed, accounting for intentionality in naturalistic terms seems hopeless within PIT: the phenomenal character of experience is metaphysically prior to the content of the experience itself. Therefore, it is the notion of phenomenal character that has to be 'translated' in a naturalistic vocabulary, in order to fit a naturalistic picture of the mind. Unfortunately, *what-it-is-likeness* is usually considered the hardest element to naturalise, given the gap between explanations at the neural or functional level and phenomenology (Levine, 2002). The unabashed naturalists of the mind like myself cannot be happy with this result.

4   The Conciliatory view: The Self-Representational Higher-Order Theory.

The two families of theories have complementary advantages and disadvantages. Naturalistic theories of intentionality have difficulties to deliver both determinate and potentially false representations. PIT can deliver such contents, but it is not naturalistic. So far, there is only one conciliatory view aiming to get the best from both: the *self-representational higher-order theory* (SHOT) by Kriegel (2011).

SHOT can be considered an intentionalist theory of phenomenal character, since the latter is ultimately nested in mental representations. However, it has to be classified within *higher-order theories of consciousness*, since it involves both first-order representations of properties in the environment and *metarepresentations* (i.e. a higher order representation) of the first-order representations. The first-order representation is constituted of a naturalistic relation defined by NERP theorists, namely informational co-variance, evolutionary history, or causation. The metarepresentation is constituted by the same natural relation, targeting the first-order representation.

Let us explain this with an example. Let us assume that Fodorean causation is the right natural relation, (the same line of reasoning can be applied, *mutatis mutandis*, to other naturalistic theories of intentionality). When I see a cow, there are two representations. The first-order content is causally determined: the first-order representation represents a cow, because the property of being a cow causes the occurrence of the first-order representation. This is in line with Fodor's semantics. The second-order content represents the first-order representation of a cow, because it is caused in turn by the occurrence of the causal relation between the cow-property and the first-order representation.

In sum, in our example, *what-it-is-like* to see a cow is to (meta)represent a first-order representation of a cow. Both the content of the first-order representation and of the metarepresentation are fixed by causal relations.[4] Moreover, my perception of a cow is intentional: it is *about/ refers to/ directed* to the cow in my environment.

Presumably, the fact that the phenomenal character occurs following the metarepresentation is what constitutes the original, intrinsic, character of intentionality. My conscious perception *of* a cow is

---

[4] The peculiarity of the Kriegel proposal is that the two representations (the first-order one and the metarepresentation) must then be *integrated* in order to generate a conscious state. This makes SHOT different from standard higher-order theories, which claim that a first-order state becomes conscious because it is the object of a higher-order state. However, for our purposes, we can skip this point.

'build up', so to say, just from an internal element of the subject: a first-order representation and a causal relation constituting the metarepresentation.[5] If this analysis really secured originality, Kriegel would succeed in having a kind of intentionality that is both original and naturalizable, since it is ultimately nested in the naturalistic relations discussed above.

However, according to Kriegel, intentionality cannot be a property of first-order representations – if the first-order representations were already intentional, then it would be like putting the cart before the horse (Kriegel, 2011: 100-101). The most fundamental kind of intentionality would have to be found in the naturalistic relations constituting of the first-order content, which has no phenomenal character on its own.

Kriegel (2011, 2013b) insists that the analogy between intentionality and ordinary representations must be treated carefully: intentional states involves representations but not all representations are genuinely intentional in his view. If first-order representations are intentional, they owe their intentionality to the (originally intentional) conscious states: at most first order representations are just *derivatively* intentional.[6]

5   Assessing the Self-Representational Higher-Order Theory

This leads us to the remark I want to make in this paper. Assuming that first-order representations are not intentional, this does not spare SHOT from solving the disjunction problem. As I pointed out in the previous section, first-order content must be potentially false in order to make sense of cases in which our experience is not veridical. As we have seen above, a never false content corresponds to a disjunctive content, whose disjuncts are all the possible cases in which that representation is recruited. Moreover, we must have a determined first-order content, because it is clear that, at least in optimal conditions, we do not experience something in between flies and little black dots when we see a fly on the table in front of us. This is also true if we assume a higher-order theory: the metarepresentation of a disjunctive first-order content ('fly V little black dot') makes us aware of 'fly or little black dot', so we become aware of everything in the environment in a disjunctive form. And this is absurd. If the first order content is indeterminate (something in between flies and little black dots), then the

---

[5] To this we should add *a unification relation* between the first-order representation and the metarepresentation (see Footnote 4) and another relation transmitting the content of the first-order state to the resulting conscious state (Kriegel 2011). I take these two to be internal relations to the subject.

[6] I think this is a fair reading of Kriegel (2011), since he holds that derivatively intentional things are so because of the attribution of intentionality by an interpreter enjoying the original intentionality. It may be the case that the interpreter interprets its first order representations to be derivatively intentional.

metarepresentation is of an indeterminate content, and therefore we always experience objects and properties in an indeterminate manner.

The worst shortcoming in accepting SHOT is that we cannot appeal to original intentionality to pinpoint the determination of first-order content and their non-disjunctive character. In Kriegel's theory, the solution of the disjunction problem is necessary to make the contents of conscious (i.e. originally intentional) mental states both determined and able to misrepresent. Appealing to original intentionality in order to solve the disjunction problem for first-order representations would involve an obvious problem of circularity: original intentionality would solve the disjunction problem for the derivatively intentional representations, whose determinacy and ability to misrepresent are necessary to make the original intentionality determinate and potentially misrepresenting in turn.

This conciliatory view relies too much on the success of NERP's theories of contents. If these theories can solve the disjunction problem, they should do it without the help of the original intentionality. Thus, the price to accept SHOT by Kriegel is to renounce one of the main motivations behind the phenomenal intentionality theory: its ability to solve the disjunction problem in virtue of the originality thesis.

If this remark is correct, then the originality thesis endorsed by SHOT adds nothing useful to NERP theories. The problems of naturalistic theories of intentionality are still there and must be solved without the help of original intentionality and there is no gain in accepting it within NERP.

PIT theorists do not gain any advantage from SHOT either. They are generally motivated by the belief that the problems of the naturalistic theories of intentionality are too difficult to solve. That is one of the reasons why they accept the originality thesis, but the attempt to pinpoint the solution of the disjunction problem onto the phenomenal character of experience brings about problems of circularity.

References

Adams, F., & Aizawa, K. (1994). "Fodorean Semantics". In Stich, S., & Warfield, T. (Eds.), *Mental Representation. A Reader.* London: Blackwell.

Brewer, B. (2011). *Perception and Its Objects.* Oxford: Oxford University Press.

Byrne, A. (2001). "Intentionalism Defended". *Philosophical Review,* 110 (2), pp. 199–240.

Dretske, F. (1981). *Knowledge and the Flow of Information.* Cambridge: MIT press.

Dretske, F. (1994). *Naturalising the Mind.* Cambridge: MIT Press.

Farkas, K. (2008). "Phenomenal Intentionality without Compromise". *The Monist,* 91 (2), pp. 273–293.

Field, H. (1978). "Mental Representation". *Erkenntnis*, 13 (1), pp. 9–61.

Fodor, J. (1990). *A Theory of Content and Other Essays.* Cambridge (MA): MIT Press.

Frege, G. (1948). "On Sense and Reference". *The Philosophical Review*, 57 (3), pp. 209–230.

Horgan, T. (2013). "Original Intentionality is Phenomenal Intentionality". *The Monist, 96* (2), pp. 232-251.

Horgan, T., & Graham, G. (2013). "Phenomenal Intentionality and Content Determinacy". In Shantz, R. (Ed.), *Prospects for Meaning.* De Gruyter.

Horgan, T., & Tienson, J. (2002). "The Intentionality of Phenomenology and the Phenomenology of Intentionality". In Chalmers, D. (Ed.), *Philosophy of Mind: Classical and Contemporary Readings.* Oxford: Oxford University Press.

Kriegel, U. (2011). *The Sources of Intentionality.* Oxford: Oxford University Press.

Kriegel, U. (2013a). "The Phenomenal Intentionality Research Programme". In Kriegel, U. (Ed.), *Phenomenal Intentionality.* Oxford: Oxford University Press.

Kriegel, U. (2013b). "Two Notions of Mental Representations". In Kriegel, U. (Ed.), *Current Controversies in Philosophy of Mind.* London: Routledge.

Kriegel, U. (Forthcoming). "Three Circles of Consciousness". In Guillot, M. & Garcia-Carpintero, M. (Eds.), *The Sense of Miness.* Oxford: Oxford University Press.

Levine, J. (2002). "Materialism and Qualia: the Explanatory Gap". In Chalmers, D. (Ed.), *Philosophy of Mind, Classical and Contemporary Readings.* Oxford: Oxford University Press.

Loar, B. (2003). "Phenomenal Intentionality as the Basis of Mental Content". In Ramberg, H. (Ed.), *Reflections and Replies.* Cambridge (MA): MIT/Bradford Books.

Lycan, W. (1996). *Consciousness and Experience.* Cambridge: MIT press.

Millikan, R. G. (1989). "Biosemantics". *The Journal of Philosophy,* 86 (6), pp. 281–297.

Searle, J. (1992). *The Rediscovery of the Mind.* Cambridge: The MIT Press.

Tye, M. (1995). *Ten Problems of Consciousness.* Cambridge: The MIT Press.

Aleksa Čupić
University of Belgrade

Relationism, Illusion and Interdependence

Abstract

In this article, I defend relationism in philosophy of perception (also known as naïve realism) from the claim that it cannot provide a valid account of misperception. I start off by presenting relationism, and a problem that is commonly associated with that theory called the problem of illusion. Afterwards, I introduce the interdependency claim, which states that perception of some qualities of an object can be interdependent. Millar (2015) argues that this claim is incompatible with the relationist's view on misperception, which states that misperception is caused by a lack of information that prompts misjudgment about a quality of a perceived object. Millar argues that there can be no loss of information in cases of illusion if the perception of qualities is interdependent. I argue against Millar by providing a weaker version of the interdependency claim that can be shown to be compatible with the relational view on misperception.

Keywords: perception, experience, relationism, interdependency, illusion.

0   Introduction

Relationism claims that perception is a two-point relation between the subject and the object of perception. One of the most common problems that relate to it is the problem of illusion. Illusions are cases where the subject perceives a quality of an object that the object itself does not instantiate. It is up to the relationist to solve this problem by taking into account the quality represented to the subject (that, again, doesn't belong to the object) in strictly relational terms. One way is to say that we misjudge the situation due to a lack of information. This article will feature a defense of this claim, which has recently received criticism from Millar (2015).

We'll start off by giving a brief outline of the relational take on perception in the first section of this article. The relational theory that will be presented in this article is a common ground theory for most relationists in philosophy of perception. After that we will introduce the problem of illusion and describe how it affects the relational view.

In the second section, we will present the interdependence claim. It states that the perception of some qualities of an object is interdependent. Call this the stronger version of the interdependency claim. It will be shown that this version of the interdependency claim cannot explain all cases of perception. The two examples that are considered in this article are cases where the subject is colour-blind or short-sighted. Because of that, we argue in favour of rejecting the stronger version of the interdependency claim, and adopting the weaker claim that will be presented at the end of the third part of the article.

In the fourth part, we consider how misperception is explained in the relational framework. Some relatonists, e.g. Fish (2009), claim that misperception happens when the subject is unaware or is lacking knowledge of some of the factors that figure in the perception of objects. Millar (2015) claims that the strong version of the interdependency claim is incompatible with the relational explanation of misperception. We will argue that his criticism cannot be applied to the weaker version of the interdependency claim, as is the case with the stronger one.

The crucial part of this article is presenting the weaker claim and showing that it does not imply the stronger version. Ultimately, we will argue for a relational take on perception which disposes of the stronger version of the interdependency claim in favour of the weaker version.

1   Relationism and illusion

Relationism (also known as naïve realism) has become popular in its ongoing struggle with the problems of illusion and hallucination, as well as in its famed rivalry with representationalism. Relationism has deserved its spotlight in philosophy of perception by becoming what Crane (2006: 128) calls one side of a chasm, created by a debate over the question of whether perception is relational or not. I think that the general outline[1] of relationism is best presented by the following claims:

i)   Perception is a relation between the subject and objects of perception and/or their qualities.

ii) The objects and/or their qualities that are being perceived are the constituents of that perceptual experience.

iii) Experience is to be understood in relational (non-representational) terms.[2]

Given that we can perceive one object in different ways, we can have a variety of phenomenal experiences of the same object. For instance, we can consider Russell's famed example where our phenomenal experience of the table changes over time due to varying ways the light is reflected upon it (Russell 2001: 2). In that situation we might perceive the table top as being brown at some times, or, we might perceive the same table top as being white at other times due to a change in the lighting in our surroundings. A similar issue arises if we consider an example where two subjects are observing the same table from two different angles. One perceives the table top as being brown and the other perceives it as being white. According to relationism, as presented above, we are then to say that because we perceive the same object in two different ways, we are in relation with two different objects, or that, in the second case, the two subjects are in relation with two different objects. This is something that relationism hopes to avoid by introducing another relatum in the subject-object relation. The third relatum is a set of standpoint conditions that is specific to each case or situation of perception. It accounts for many different conditions that might affect the phenomenology of perception, such as: space-time positioning, distance, clarity of view,

---

[1] For a more substantial expansion on the three claims see: French (2014), Chirimuuta (2015: 43-49), Campbell & Cassam (2014: II).

[2] Relational terms are those that refer to the represented object or quality without including the subjective character of experience.

arrangement of perceived objects, lighting, contrast, etc. We are then to say that, in the first case of the example presented above, we indeed are in a relation with the same object, the only difference is that the standpoint conditions changed over time. Because of that the table appears different but it is still the same object of perception. As for the second case, we are to say that the two subjects are in relation with the same object, but are observing it from different standpoint conditions, and thus the table only appears different to them.

It should also be noted that relationism comes in a variety of different versions. This article will not be concerned with a specific version of relationism - when referring to relationism, the kind of theory that we have in mind is the one presented above. We believe that most relationists would agree with what has been stated so far.

Now that we've outlined relationism, we can turn our attention towards illusions. An illusion is a kind of experience where an object is perceived as being different than it really is. To put it more precisely, an illusion[3] is a situation when:

i)   S perceives O as being F.
ii) F is not actually instantiated by O.
iii) G is instantiated by O.
iv) S perceives F instead of G.
v) (iv) is because of circumstances C that S is unaware of. (Optional)

Consider some examples of such situations. Example 1: S perceives a blue wall when the wall is actually white but appears blue because of blue light reflected on the wall, which comes from a light source that S is unaware of. Example 2: S perceives a white rectangular tower in the far distance. The tower is actually round but only appears to be rectangular to S because of the distance between them. Example 3: S is looking at a circle on a piece of paper, but what he isn't aware of is that a distorting lens was placed in front of his eyes, so what he actually sees when looking at a circle is an ellipse.

---

[3] For a more detailed account of illusions see Smith (2002: 22-64).

Now, why are illusions considered a threat to relationism? If perception is taken to be a relation between the subject and the real[4] quality of a (mind independent) object, then, how are we to explain the additional qualities, represented to S in examples 1-3, which are actually not instantiated by objects that S is perceiving? I maintain, as it will be examined later, that most relationist tend to fall back on the standpoint conditions when it comes to answering this question. We can now move on to examine the interdependency claim.

2   Interdependence

"The distinct phenomenal features instantiated by a given perceptual experience are, to a significant extent, interdependent." (Millar, 2015: 612) Let us first consider what "interdependent" means. Millar doesn't give a rather clear definition of it, so we'll try to give one based on the following quote: "when you see the shape of a colored object, you see the object's shape in virtue of seeing its color." (Millar, 2015: 613)

If we can only see the object's shape in virtue of seeing its colour, then the perception of the shape of an object is dependent upon the perception of its colour. Thus, to say that perceiving A is dependent on perceiving B means that: we perceive A in virtue of perceiving B.[5] Let's examine this with the help of an example. Imagine seeing a coloured red square drawn without an outline on a blank, white paper. It is obvious that by perceiving that which is red on our background we perceive a square. For if the square was to be completely white, it would blend in perfectly with the background, making it impossible to be perceived. And if we cannot perceive the distinct colour of an object (as contrast to a background in this case), then it seems obvious that there is no shape to be perceived as well. Simply put, take away the colour and you take away the shape as well. So far, it seems plausible to say that in perception, shape is dependent on colour.

The common meaning of "interdependent" is "dependent on each other". So, if we were to say that A and B are interdependent, that would mean that A is dependent upon B and that at the same time

---

[4] The reason why I use the term "real" here is to differentiate between real and illusionary qualities of objects in experience. The former are the perceived qualities that are instantiated by the object (such as quality "G" in the argument presented above). The latter are the perceived qualities that are not instantiated by the object (such as quality "F" from the argument presented above).

[5] We concede that it is not very clear how we are supposed to understand "in virtue of". To make things clearer let us say that to perceive A in virtue of perceiving B is to say that when we perceive A, we immediately perceive B.

B is dependent upon A. We already showed how shape can be considered as being dependent on colour; now we ask ourselves if it is possible the other way around? Can it be shown that colour is dependent on shape? In that case, we would need to give an example where, when perceiving an object, we perceive the object's colour in virtue of perceiving its shape. Let's go back to the example with the red square. Imagine the same situation as before. Are we entitled to say, similar to what we said about colour, that by perceiving that which is a square on our background we perceive that which is red? For now, it seems that we can. We can thus move on to conclude that it seems plausible that, in perception, colour is dependent on shape.

Now imagine a third case. Say a colour-blind person takes a look at our red square drawing. He is obviously looking at the same object as we are. Both objects of his and our perception (the image on the paper) are the same, the difference is in how we perceive it. Where I see a red coloured square, he sees a grey coloured square. As we see it, the only way by which we could include this issue into the relational view on perception is to say that the difference in the perception of the two mentioned subjects is to be explained by the difference in the set of standpoint conditions. His set of standpoint conditions contains colour-blindness, ours does not. But does the interdependency claim apply to his perception as well as ours? More precisely we can ask: does he perceive the object's shape in virtue of perceiving its colour in the same way as we do? Furthermore, does he perceive the object's colour in virtue of perceiving its shape? Even though an extreme relationist would probably pressure us to say yes on the basis of standpoint conditions, as he would like to make a relational explanation of both our perception and that of a colour-blind person possible, we are inclined to say no. Because clearly, the person from the third case is not perceiving the colour of the square. However, he does perceive a different colour even though it is not the one instantiated by the object. Yet, he can clearly, without any problem, perceive a square. In conclusion, he does not perceive the shape of the object in virtue of the colour that the object instantiates, since he does not perceive the actual colour of the object. Furthermore, he cannot perceive the colour of the object in virtue of perceiving its shape for the same reason. But he can, however, perceive the shape of that object perfectly and without any problems. And it also seems that he is perceiving shapes in virtue of colours (and vice versa), the problem is that the colour at hand isn't the colour instantiated by the object.

Now given the nature of this case, we could go back and try to fix the dependency definition: to say that perceiving A is dependent on perceiving B is to say that we perceive A in virtue of perceiving B. Instead of that we could say: to say that perceiving A is dependent on perceiving B means to say that we perceive the object as having (a kind of) A in virtue of perceiving that the object has (a kind of) B. Now this seems like a weaker claim. It also seems that it can be applied to our perception and the perception of the colour-blind person, because he perceives the shape of the object in virtue of it having (some) colour.[6]

I believe that the second version of the definition of dependence can also work well for our fourth case, which would feature a person that has a different kind of sight impairment, which allows him to perceive colour exactly as it is, but it does not allow him to perceive shape correctly. Such problems are usually the result of bad eyesight (e.g. short-sightedness, double vision, etc.).

3   Judgement and Millar's objection

A common way of explaining misperception in the relationist framework is to blame judgement. (Millar 2015: 213-216) In that case we would say that the illusionary experience is characterized by a lack of knowledge or awareness as to what the illusionary factors of our experience are. I call them illusionary factors because these factors are what makes our experience an illusion. If those factors were to be removed, we would have no illusion in our experience. The lack of knowledge of such factors is what prompts misjudgement in the case of illusion. So instead of saying that we misperceived an object, we say that we misjudged it on the account of having no knowledge of the illusionary factors present. One example of such illusionary factors is the distorting lens from Fish's example (Fish 2009: 163). For example, imagine a situation where a subject is looking at a circle trough a distorting lens that he isn't aware of; so instead of seeing a circular shape he actually sees an elliptical shape. Here we have an example where the "blaming judgment" can be applied. Fish would like us to believe that because of the lack of knowledge or awareness of the distorting lens (and the lack of knowledge of the shape of the object that follows), the subject is misjudging the shape of the given object. Another example of this issue might be given by the situation that

---

[6] We do not need to specify the colour, but for the sake of the argument, let's make it clear that the colour-blind person is not perceiving the same, exact colour as the colour of the square.

was examined previously. Namely, where the subject observes a tower in the far distance. There is no lack of knowledge when it comes to this situation, other than that of shape, since, unlike the distorting lens, the distance, which plays a role of the illusionary factor in this example, is clearly apparent to the subject. Fish might push forward with this example by saying that the thing that the subject is not aware of is not the distance itself, but rather the role that it plays in constituting his phenomenological experience of the tower. And it is this lack of awareness that leads him to judge that the tower's shape is rectangular instead of round.

Millar (2015) states that the lack of knowledge that prompts misjudgement in such a case cannot explain the illusionary experience of the subject from a relational point of view. If we were to explain the illusion in this situation we would need to mention that the elliptical shape is part of the phenomenology of the subject. That is the key point of the illusion. Now if we were to adopt the relational view and the previously examined (strong version of the) interdependency claim, we start to notice some problems with this solution. If the subject is perceiving the colour correctly, then, he cannot misperceive the shape (because of the interdependency of perceiving colour and perceiving shape), which means that there can be no elliptical shape in the phenomenology of his experience. And if the subject cannot misperceive the object's shape it follows that there can be no lack of knowledge or awareness of the shape. This in turn means that there can be no misjudgement, since there was no misperception in the first place.

We claim that this issue can be avoided by disposing the strong version of the interdependency claim, and adopting the weaker version instead. If we consider Fish's example one more time, we can say that the subject perceives the object as having a kind of shape in virtue of having a kind of colour.[7] The weaker claim does not prevent the subject from misjudging the situation since it only claims that the fact that we are perceiving a shape is due to the fact that we are perceiving colour and vice versa. It does not specify, precisely, what the colour or shape is. The difference that is introduced here is the difference between saying that an object has colour or shape and saying that an object is of a specific colour or shape. The former (which is the weaker interdependency claim) does not imply the latter (while the stronger claim does). For example, if someone says that a square has a kind of colour, it does not imply that the colour is red, or, if someone says that a red

---

[7] I believe that using "certain" instead of "kind of" would be more precise, but it would misconstrue the application of the definition.

object has a certain shape, it does not imply that the shape is square. By adopting the weaker instead of the stronger claim, there is nothing that prevents the subject from making the misjudgement that was presented above, because the implication mentioned cannot be made.

4   Conclusion

We started off with a general outline of the relational take on perception. We noticed that perception is explained as a three-point relation between the subject, the object and the set of standpoint conditions. We then presented the strong version of the interdependency claim and noticed that it cannot be applied to all cases of perception. The cases that were presented in this article are that of the colour-blind person and a short-sighted person. It was shown that the first case is incompatible with the claim that the colour of an object is perceived in virtue of the shape of an object and vice versa for the second case. To address this issue, we introduced a weaker claim that works for all cases mentioned. The advantage of the weaker version is that it does not imply the stronger version, and, because of that it does not prevent us from making the claim that misperception happens because of the lack of information that prompts misjudgement. We conclude now that, for all cases mentioned, relationism works well with the weaker claim and that it evades Millar's criticism of Fish's explanation of misperception.

References

Campbell, J. & Cassam, Q. (2014). *Berkeley's Puzzle: What does experience teach us?".* Oxford: Oxford University Press.

Chirimuuta, M. (2015). *Outside Color Perceptual Science and the Puzzle of Color in Philosophy*. Cambridge: MIT Press.

Crane, T. (2006). "Is there a perceptual relation?". In Gendler T. S. & Hawthorne J. (Eds.), *Perceptual Experience*, Oxford: Clarendon press pp. 126–146.

Fish, W. (2009). *Perception, Hallucination and Illusion*. Oxford: Oxford University Press.

French, C. (2014). "Naïve realist perspectives on seeing blurrily". *Ratio*, 27, 4, pp. 393-413. doi:10.1111/rati.12079

Millar, B. (2015). "Naïve Realism and Illusion". *Ergo,* 24 (2) pp. 607–625.

Russell, B. (2001). *The Problems of Philosophy*. Oxford: Oxford University Press.

Smith, A. D. (2002). *The problem of perception*. Cambridge: Harvard University Press.

Niko Šetar

University of Maribor

To Be or Not to Be – Quality of Life Between Antinatalism and the Repugnant Conclusion

Abstract

In his 2006 book, titled *Better Never to Have Been: The Harm of Coming into Existence*, David Benatar defends the controversial theory of antinatalism, i.e. the notion that no more sentient beings (humans) should be brought into existence. Benatar's position is composed of two main elements. According to the first, by bringing a person into existence, we are responsible for everything bad that might befall that person, while we may take no credit for what good they might experience. The second element states that after examining all bad and good in all individual lives (the amount, intensity, etc.), we would be bound to conclude that most of all human lives are, as a whole, bad and not worth living. It has often been claimed that upon accepting the second element, we have to acknowledge that not only should we no longer produce new people, we should actively get rid of currently existing people; namely, commit an extinction-scale suicide. While this is an interesting possible objection, this article will focus on the first element and the argumentation scheme that supports it, along with possible consequences of refuting the argument.

The argument for the first Benetar's element revolves around an asymmetry between absence and presence of pain and pleasure. Benatar claims that, while presence of pleasure is always good, and presence of pain is always bad, this does not hold true about the absence of the two. He maintains that "absence of pain is always good, even if that good is not enjoyed by anyone" and "absence of pleasure is not bad unless there is somebody for whom this absence is a deprivation." (Benatar, 2006) For any potential person X, this may be summarised, if awarded arbitrary values of 1 for good, 0 for neutral, and -1 for bad, in the matrix below:

	X exists	X does not exist
Pain (present/absent)	Bad (-1)	Good (1)
Pleasure (present/absent)	Good (1)	Neutral (0)
Sum	Neutral (0)	Better than neutral (0.5)

This does seem plausible and is thus supported by several additional arguments. First of these is the view that, even intuitively, we seem to have a sense of moral duty not to bring people suffering into existence, while we have no sense of duty to bring happy people into existence. Benatar argues that this supports the asymmetry, in that the asymmetry best explains why we have these notions. He alleges that they are based on the passive fact that we do not perceive the lack of good that comes from not giving life to a person as deprivation, while we do perceive the lack of bad (pain) that comes from not giving life as a benefit. This raises the question of who is deprived or benefited in this scenario, but we shall avoid addressing the non-identity problem for now. Relatedly, potential interests of a child are generally used to justify why one should not have a child, yet never (according to Benatar) to justify why one should have a child, exactly because of the abovementioned notion of a duty not to cause suffering but no the notion of a duty to create benefit. Yet, as we will attempt to establish later, duty to create benefit, albeit weaker in nature than duty to avoid suffering, can still be identified.

The third argument supporting the asymmetry Benatar uses is that while we can regret creating a suffering child, we cannot regret not creating a happy child. For Benatar, it is a matter of rational impossibility. This is because our regret upon creating a suffering child has a direct object – namely, that child – yet our hypothetical regret about not having created a happy child has no object, since the child had never come into existence. The last of the four arguments refers to

distant suffering and unpopulated areas. Most of us do, in fact, feel sad about the suffering of, for example, starving people in Africa. However, nobody feels sad about an uninhabited island because there are no people on it to enjoy its benefits. Similarly, argues Benatar, we do not regret Mars being unpopulated, but we would likely mourn the suffering of Martians if Mars was populated and Martians were indeed suffering.

These arguments seem as sound and plausible as the original formulation of the asymmetry. Nevertheless, something about them is so deeply counterintuitive that they merit a closer examination and a justification of scepticism that arises about them.

Firstly, let us examine the last argument supporting asymmetry. It is immediately noticeable that the example of Mars is entirely redundant, as it is an uninhabitable area that, even if inhabited, does not have the capacity to provide means for a good life. Given its lack of resources and other liveable conditions, every being on Mars would be subject to a nasty, brutish, short life, which would indeed not be worth living by even the most generous standards. For Benatar's asymmetry on this matter to hold true, we have to consider a more promising uninhabited environment, one with plentiful resources and favourable conditions. Imagine, for instance, a deserted island with food galore, drinkable water, all necessary ores and minerals, as well as a temperate climate. Visiting that island, it would be extremely easy to conceive of a possible world or a possible future where the island is occupied with people living lives of great prosperity and, thus, pleasure. Plausibly, one could also regret the island not being populated by people enjoying that pleasure without identifying those people or presuming that they are currently living.

Next, we will look at the third argument supporting the asymmetry, which claims we can regret creating a suffering child, but cannot regret not creating a happy child. The first issue one needs to point out is the reliance of this asymmetry on retrospection. We can obviously imagine grief-struck parents of a severely handicapped child whose (likely very short) life contained nearly nothing but suffering. The issue is that if the child's life was at least of average quality, they would not regret its conception. Even dismissing the fact that their grief might have only been misinterpreted as regret, it still holds that they are regretting the child's suffering as a consequence of her conception and not her conception *per se*. Analogously, we can easily conceive of a reasonably well-off couple in, say, their fifties, who are reflecting on the benefits they could have bestowed upon their children, had they had them. They, again, regret not conceiving a child not because of her non-

conception in itself, but because of the good of which they had deprived their possible child. Should this be false, we would have to acknowledge regret is of and because of the conception (or existence) of a person, and not remotely related to the quality of her life, rendering the pain-pleasure content of her life irrelevant. Furthermore, we can regret not benefiting our non-existing children for the same reasons we can think of children's potential interests before their conception: when we imagine a child, we ascribe her a partial identity, even if that partial identity only consists of the pain/pleasure value of her life. This does not, as one might object, lead to Saul Smilansky's (2013) dilemma of choosing between regretting every person's existence or not regretting any person's existence at all. The premises that lead to this dilemma are far too impersonal (as are, as we will see later, Parfit's premises in the Repugnant Conclusion). Rather, it leads to a more personal conclusion, where it is possible to rationally both regret a suffering person's existence and also regret a happy person's non-existence.

This brings us to Benatar's second and first supporting arguments for the asymmetry, the two of which I shall consider as mutually reliant and attempt to argue against both on the same grounds. Firstly, potential suffering is more determinable than potential happiness. We have technologies that enable us to find out whether the embryo will develop certain genetic defects that might cause it to suffer, and even the probability of not immediately detectable genetic issues that might cause suffering may be estimated by examining, e.g. family history. However, there are no genetic indicators of increased happiness which would allow us to say, with any degree of certainty, that a person will be predominantly happy in her life. Yet, what if there were such indicators? Suppose scientists find a 'happiness gene', which either increases overall happiness or decreases sensitivity to life's toils, increasing happiness consequentially. Would we not, if that were possible, ascribe this happiness (in the sense of partial identity hypothesised in the above paragraph) to our child in foresight, just like we do potential suffering-inducing deformities? We certainly would, which would plausibly lead us to a graspable notion of a duty to create happy people.

We can now return to the main asymmetry and impose the conclusions above on it. If "absence of pain is always good, even if that good is not enjoyed by anyone" and "absence of pleasure is not bad unless there is somebody for whom this absence is a deprivation", then the absence of pain is good because we have a determined pain in mind when we are talking about potential pain of a (yet) non-existing child, which may be attributed to that child in a more concrete fashion than its

potential pleasure, which is less determinable and therefore more abstract. The asymmetry is not only counterintuitive, but also stems from an *ad ignorantiam* about the nature and determinacy of pleasure. If both potential pains and potential pleasures were determinable to the same degree, either by science or otherwise, both would serve as reasons to have or not to have children, and both would be considered equally relevant in all asymmetries, rendering them symmetrical; that is to say, absence of pleasure would be considered a deprivation just as absence of pain is considered a benefit in Benatar's argument.

However, does that mean that we now have a moral duty to create a happy child? We are faced with this question in the Happy Child Dilemma. While we have to consider, as Benatar also points out in one of the alternatives to the asymmetry he analyses, a certain impairment on pleasure of the parents of the child (e.g. health risk to mother, financial or material risk to both), we may presume, even without acknowledging any intrinsic value of life, that the lives of parents would preserve their prior quality due to their satisfaction of raising a happy child, while the life of a happy child adds to the overall happiness in the world.

What happens if we accept all this and answer the above question affirmatively leads us to "The Repugnant Conclusion", a chapter in Derek Parfit's book *Reasons and Persons.* If a moral obligation for having children is imposed, we are facing imminent population growth. According to Parfit, population growth has both positive and negative effects. The positive ones are generally transitory effects, such as a sudden increase in purchasing power and economic growth that follows. The negative ones are cumulative effects, such as the distribution of finite natural resources among an increasing amount of people. These effects manifest differently, depending on how the population increases. Parfit calls for two possible manners of this increase: *Growth* and *Replacement.* In the Growth scenario, the increase in population is greater than 1 child per 1 parent. Due to transitory effects, Growth is a good strategy for the first couple of generations, but soon leads to cumulative effects that start creating a negative impact which does not improve anymore. In the Replacement scenario, where the increase is exactly 1 child per parent, the population remains the same, creating no transitory or cumulative effects, maintaining the status quo. A third option is a combination of both – Growth for the first few generations, switching to Replacement as cumulative consequences begin to appear. However, Parfit claims that this leads to a cataclysmically sudden loss of transitory effects, which lead to worse negative consequences than

gradually losing benefits and resources in Growth scenario. The former is, therefore, the most favourable.

The question that arises now is the speed of growth. Parfit proposes two populations, A and B, which start out with equal populations, with A growing moderately and B rapidly. After 300 years, the population of B is twice as large as population A, meaning that the quality of life is somewhat less than twice lower in B than in A – the initial quality of life was sufficient for lives in both populations to still be well worth living. This brings forward more questions. When will the population of B exceed the resource distribution threshold when lives cease to be worth living? Is population B better than A, or, to paraphrase, is the amount of lives lived more relevant than the quality of each individual life?

One way to work through this question is to follow the *Impersonal Average Principle*, which states that the best outcome is the one where people fare best on average, that is to say, where the average quality of individual lives is the highest, with the highest amount of average V (unit of whatever value makes life worth living) per person. However, this is not without its disadvantages, one of the major ones is that the death of people with the greatest amount of V decreases everyone else's average, or that a 'better' world might well be secured by killing off those with the lowest amounts of V to increase the average.

Parfit then proposes that loss of life quality might be justifiable if the quantity of lives lived is sufficiently large. But when does this come to pass? We can use two instrumental questions to attempt to answer that. Firstly, if people have lower quality of life on average, can, under any circumstances, a higher number of those people outweigh an individual's lack of life quality? If yes, what are the relative values of quantity and quality, and are they fully reciprocal? If again we answer affirmatively, we arrive at the *Impersonal Total Principle,* where the total sum of all V in both populations must be considered to determine which one is better. As we said, if the population of B is twice as large as population A, the average amount of V in B is marginally larger than half of the amount of V in A (for example, in V(A) = 10 and in V(B) = 5.000001) then the total amount of all V in B must be marginally larger than the total in A (for example, if population of A = 1,000,000 and population of B = 2,000,000; the total of V in A = 10,000,000 and the total of V in B = 10,000,002), making B the better of the two populations.

Here is where we start to slide down a logically justified slippery slope into the Repugnant Conclusion. If population B is indeed better than population A, it must also be true that a population C, which grows at twice the rate of B – four times the rate of A – is better than both populations A and B, despite the four times smaller amount of V (as long as there is *any* amount of V, life is still worth living). Once we have accepted this, we must also accept that there must exist a population Z, which is, in accordance with the English alphabet, 2^{26}-times larger than the population of A, with each individual in that population posessing 2^{26}-times less V than in A. That means lives in Z are barely worth living, yet still worth living. Between this and the Impersonal Total Principle, we must accept that no matter how (relatively) awful lives in Z may be, Z is still a better alternative to A.

Evidently, this would be best avoided, if possible. And possible it is. Douglas Portmore (in Portmore, 1999) argues that the Repugnant Conclusion does not follow from the Impersonal Total Principle. Portmore claims that there are three possible ways Z can be barely worth living: *Drab Z, Rollercoaster Z* and *Short Z.* In Drab Z, lives are dull with virtually no pain, but also virtually no pleasure. In Rollercoaster Z, they are filled with ecstatic pleasures followed by equally intense and frequent agonies. In Short Z, they are of the same quality as in A, but extremely short. Parfit describes Drab Z and Rollercoaster Z in *Reasons and Persons*, but in Portmore's view, they still do not have any repugnant implications due to the discontinuity view, which proposes that the V we might find in A is not the same as in Z. For example, Z contains a certain V1, while A also contains a V2, which makes life not just barely, but wholly worth living. Because the relation between V1 and V2 is discontinued, there does not need to be an amount of V1 for any amount of V2. In fact, even a marginal amount of V2 outweighs any amount of V1. This makes Z worse than A regardless of how the two worlds compare in terms of V1. Parfit himself states that any one person would likely prefer to live in a Century of Ecstasy (100 years of constant euphoria) than in an eternity of Drab Z, which confirms the idea of discontinuity. In this context, B may still be preferable to A, but the descent down the slippery slope is cut short long before reaching Z.

Let's now examine more closely what is at stake in the choice between the *Century of Ecstasy* and an eternity spent in Rollercoaster Z. In the latter, a person will experience, even though rivalled by agonies, all the pleasures they would in a Century of Ecstasy. This leads to "Well-being and Time" by David Velleman, where the author describes a life beginning with a difficult childhood and

youth, followed by a successful middle-age and a happy retirement, as opposed to a life that starts with a wonderful youth, followed by a series of personal disasters and a sorrowful, regretful retirement. A Global View on the matter shows that an improving life is better than a decaying one, as success and happiness attribute meaning to earlier pains and troubles, while subsequent 'disasters' nullify earlier pleasures. Thus, we can imply that a Century of Ecstasy is better than Rollercoaster Z, as every pleasure experienced in the latter is nullified by the agony that follows. The discontinuity is also evident here, as no ecstatic or agonizing V can be proportional to a comparatively 'drab' V we can find in A.

All that is left is Short Z, for which it would seem that it is not repugnant at all. A life of a child leading a fulfilled and happy life that is abruptly ended in a car accident is, by all means, tragic, but nothing about that life itself, before its end, can be classified as repugnant.

Conclusion

The conjunction of the Impersonal Total Principle and the Discontinuity View therefore shows that the extreme population of Z cannot be reached before the chain snaps at a point where life stops being worth living. This seems to imply that when faced with the Happy Child Dilemma, we have a duty to create a child, but only insofar as we can be convinced that this child will be happy in a sense that her life will be reasonably worth living. However, what is left do determine is when does life stop being worth living? What are the criteria that need to be established for a person to foresee the quality of their child's life and decide whether or not it will be worth living? Taking into account what we have described while replying to Benatar's asymmetries, pleasures and pains, and implicitly the general quality and worth-livingness of life, are too undeterminable for a nominal universal criterion. Therefore, it may be concluded that neither pains nor pleasures are sufficient reasons for not bringing or brining a person into existence, respectively, and that *pro tanto* reasons for and against conception are also insufficient because of the very fact that they fail to consider contingent value of life in the context of pain/pleasure content. It may, however, be concluded that if we do not have a duty to bring someone into existence, then neither do we have a duty not to bring them into existence. What we do have is the right to bring them into existence

as long as we can, with what is determinable before birth or even conception, reasonably support the claim that our child's life will be not necessarily happy, but worth living.

References

Benatar, D. (2006). *Better Never To Have Been: The Harm of Coming into Existence.* Oxford: Clarendon Press.

Parfit, D. (1986). *Reasons and Persons.* Oxford: Clarendon Press.

Portmore, D. W. "Does the Total Principle Have Any Repugnant Implications?" *Ratio,* 12 (1), pp. 80–98. doi:10.1111/1467-9329.00078

Smilansky, S. (2013). "Morally, should we prefer never to have existed?" *Australian Journal of Philosophy*, 91 (4), pp. 655–666.

Velleman, J. D. (1991). "Well-Being and Time". *Pacific Philosophical Quarterly*, 72 (1), pp. 48–77.

Jelena Pavličić

University of Belgrade

Scepticism and Contextualist Diagnoses of the Problem

Abstract

Epistemic contextualism, which encompasses conversational and inferential contextualism, is one of the most widely discussed epistemological projects nowadays. However, the bulk of the literature on epistemic contextualism focuses on the analysis of the advantages and difficulties posed by these types of theories, predominantly treating them separately. The subject of this paper is the evaluation of these theses in light of each other. First, we provide an overview of the compelling contextualist diagnosis of the problem of scepticism - one that accommodates both sceptical and anti-sceptical intuitions while preserving the highly intuitive principle of the closure of knowledge. Both versions of epistemic contextualism share this core. However, there exist several differences between the two camps on the way to reach this diagnosis, one of which is discussed here - the explanation of how we fell into the sceptical trap in the first place. Conversational contextualism strives to explain this by adopting an error theory that has been the target of many objections that can be lined up into two groups: those that aim to show that (i) error theory is not plausible, and (ii) the contextualist appeal to error theory leads to specific problems within the contextualist approach itself. Presenting one argument of the second kind (Schiffer's argument), we come to its important role in forming objections that, it turns out, conversational contextualism fails to overcome. Thereafter we outline an inferential diagnosis of scepticism which is not committed to the highly problematic error theory and indicate one of its (dialectically observed) most significant aspects - shifting of the burden of theory. We further argue that these results contribute to the notion of inferential contextualism as a favoured resolution of the problem of scepticism.

0   Introductory remarks

Contextualism in epistemology is the position that the truth condition of knowledge sentences may vary according to the standards at play in a particular context. Perhaps the main advantage of this proposal in the theory of knowledge is that it is allegedly able to overcome the longstanding problem of philosophical scepticism in, as we are about to see, a refined and effortless manner. However, there exist major discrepancies between two of the main contextualist camps: conversational (David Lewis, Stewart Cohen, Keith DeRose and others) and inferential (Michael Williams), which differ with respect to the manner by which they claim to reach this solution. We will focus on one aspect in which the two solutions differ: their answer to the question of how we ended up with the problem of philosophical scepticism in the first place. The main contribution of this paper is that it shows how the comparative effectiveness of these two anti-sceptical strategies crucially depends on their answer to this question. The paper proceeds as follows.

First we introduce the sceptical problem (§1). Then we provide an overview of the anti-sceptical strategy of the "conversational" contextualism (§2). We show that in spite of its many advantages, this line of argumentation faces criticism that is very difficult for the contextualist to counter. In (§3) we propose another contextualist solution that is endorsed by Michael Williams and compare these two theories.

1   The Sceptical Problem

The philosophical problem that we might have no knowledge about the world around us is known as the problem of *scepticism about the external world*. The problem arises due to the "sceptical hypotheses" such as that we might currently be dreaming; that we are subject to some massive deception (Matrix); that we might be a brain in a vat (that is a disembodied brain connected to electrodes that renders perceptual experiences identical to those we are having right now), etc. These hypotheses are "sceptical" because of their two main characteristics: (i) their truth is incompatible with every proposition about the external world such as we commonly take to be a subject of knowledge. In simple terms, this means that if one of these hypotheses were true, then all our current beliefs about the external world would be false. And (ii) they are perfectly compatible with all the evidence in favour of those common beliefs. That is, if the deception was good enough (and let us assume that it is), then we (e.g. BIV) would experience a perfectly realistic

virtual world in which we are discussing sceptical possibilities, reading these very lines, and so on. That is to say, there is nothing on the phenomenological level that we can pick out that can neutralize these possibilities. And for the sceptic, these insights are quite sufficient to argue as follows:

1. I don't know that I am not a handless brain in a vat.

2. If I don't know that I am not a handless brain in a vat, then I don't know that I have hands.

3. Therefore, I don't know that I have hands.

This is *the sceptic's argument*. Premise (1) is based on the insight that, however improbable it seems to assume that we are a BIV, it is reasonably clear that we do not know that we are not. As for premise (2), which at first sight may appear weaker than the first, it encompasses the epistemic principle that knowledge is closed under a known entailment, a view which is held as extremely plausible by many philosophers. The principle is as follows: if we know something and we know that that thing entails another thing, then we know that other thing. For example: we know that at this very moment our bicycle is in the garage, and we know that if our bicycle is in the garage then it is not somewhere else, for example at the flea market. Hence, we know that our bicycle is not at the flea market. Rejecting this principle means allowing something what Keith DeRose labels "abominable conjunctions" (1995: 27-9). Namely, it would permit us to say things like: (1) "I know that my bicycle is in the garage, but I don't know whether it has been taken to the flea market" or (2) "I know that I have hands, but I don't know that I'm not a (handless) brain in a vat."

With this in mind, we may say that the paradox lies in the fact that the sceptic's argument is valid and uses premises that seem true. On the other hand, it also seems that the denial of the conclusion of the argument is true. Thus, evidently, these three propositions cannot all be true:

(1) I know that I have hands.

(2) I don't know that I am not a BIV.

(3) If I know that I have hands, then I know that I am not a BIV.

Something plausible, (1), (2), or (3), has to be discarded, leading to questions which are far from straightforward – "Which of these statements is to be rejected?" and "How do we go about rejecting a plausible statement?" Epistemologists have made many attempts to solve the puzzle by

trying to negate proposition (2) or (3). But, as the contextualist emphasises, it appears utterly arbitrary merely to reject one of them on the basis of the other two (Cohen, 1988: 94). For this reason, he proposes a different solution, one that can meet (1) our strong conviction that we have knowledge of the external world and at the same time explain (2) scepticism's intuitive force, while preserving (3) the highly plausible closure principle.[1] Both versions of contextualism share this core. In what follows, we will explore how contextualists attempt to achieve (1), (2) and (3), in order to compare the efficacy of these two anti-sceptical strategies. We begin with the argument of conversational contextualism.

2   Conversational Thesis

Proponents of conversational contextualism start from the observation that numerous words that we use in ordinary language are context-sensitive. Consider the word "flat". Most of the time when we say for some table that it is flat, what we mean is that it is sufficiently flat for our ordinary purposes; to wit, it is suitable for dining, writing, and so on. However, in a physicist's context, for the purpose of a scientific experiment, it may be that that very same table simply isn't flat enough. As Peter Unger (Unger, 1975) and later Stewart Cohen (2005: 60) pointed out, one may lead this strategy to the extreme and introduce the context where even microscopic irregularities count, such that in that context we can truthfully say "nothing is really flat". Does that mean that our everyday claim "the table is flat" is not true? That is not the case, since according to the widely accepted scalar analyses of gradable adjectives (see, for example, Kennedy, 1999; Blome-Tillman, 2008) the two sentences "the table is flat" and "the table is not flat" are not actually incompatible due to do the fact that they express a different proposition in a different context of utterance and thus have different truth-conditions – simply put, they basically mean different things. This inspection is in alignment with the view that the standards for the application of the adjective "flat" are context sensitive. Contextualism maintains that the manner in which we use the verb "know" is akin to that of the adjective "flat". The following example should serve to illustrate the role of the context in usage of the verb "know" in ordinary situations.

---

[1] Then again, there is an exception among contextualists regarding (3). Namely, contextualist Mark Heller is one of the few authors that provides a theory that *rejects* the *closure* principle (Heller *Mark. 1999.* Relevant Alternatives and Closure. *Australasian Journal* of *Philosophy*, *77(2): pp. 196–208).*

Hana is in Maribor with a group of friends. They are about to go rafting in Soča valley, and some of them, who've never been to Ljubljana, insist on visiting the city beforehand. After a long and exhausting deliberation they agree to take the bus to Ljubljana later that evening. Trying to check the timetables, they realize that they have no internet access, at which point Hana says: "Guys, don't worry, I know that a bus leaves to Ljubljana five minutes before midnight because I checked the timetable last night. I'm going to my room now and see you later." Since nothing really important is at stake one of the others says: "OK, Hana knows the bus leaves five minutes before midnight, let's make use of the day and leave for Ljubljana at that time." The group agrees and makes way for the city centre.

It is evident that Hana's friend made a knowledge attribution and that the rest of the group tended to agree that the sentence is true. And as far as contextualism is concerned, this claim is true. We ascribe knowledge to people on the basis of this type of testimony all the time because the testimony meets the standards for attributing knowledge in everyday situations. But let us now shift to another context that evokes somewhat different standards.

In the same setting, consider an alternative scenario: Hana is in Maribor with the same group and they want to know at what time the bus leaves. Hana informs them of the schedule she saw and then leaves to her room. But then, one of the friends points out that she has promised to meet a business partner in Ljubljana early in the morning; the meeting is important and she cannot afford to risk it. She says "Maybe the schedule that Hana saw was the local bus to *Ljubljanska street* and not the bus to Ljubljana. Maybe the bus will not show up at the specified time and, being the last bus, we will have to wait until the following morning". Given the importance of arriving on time, Hana's friends start to consider their previous knowledge attribution and admit that Hana doesn't really know that the bus leaves five minutes before midnight.

We have just seen that the standards for knowledge attribution have changed between the two examples. These observations lead us to ask the question: which standard is correct? Contextualists hold that neither standard is simply correct or simply incorrect. Rather, it is the context that determines which standard is correct (Cohen, 1999: 59). Standards vary according to the features of the conversational context. As a result of this variation in epistemic standards, the contextualist would consider it possible that, when two speakers, speaking in separate contexts, simultaneously say "*Hana knows* that *p*" one of the statements is true and the other false, even though both are

speaking about the same individual (Hana) and the same proposition (the bus leaves five minutes before midnight).

Before we revisit the problem of scepticism, and examine its interpretation in light of the contextulaist thesis, let us look at one point of *significance* for this argumentation. "Context" among conversational contextualists is entirely understood in terms of conversational factors like purposes, interests, presuppositions, intentions, etc. However, conversational factors that are relevant in shaping the standards for knowledge are not the factors of the subject of the knowledge ascription (Hana's intentions) but rather, the conversational factors of the attributer's situation (that is the intentions and purposes of Hana's friends). Here is Cohen:

> "Now, when I say 'contexts', I mean 'contexts of ascription'. So the truth- value of a sentence containing the knowledge predicate can vary depending on things like the purposes, intentions, expectations, presuppositions, etc., of the speakers who utter these sentences". (Cohen, 1999: 57)

Keeping this important feature of conversational contextualism in mind, let us return to scepticism. We are currently in the context of philosophy in which, presumably, we tend to agree that the sceptical argument is irresistible, that the BIV hypothesis is difficult to rule out, etc. Nonetheless, when we look back to our everyday practice we recognize that we are not troubled by the problem of philosophical scepticism. In ordinary epistemic contexts, the standards are usually much more lenient and easy to fulfil, and we tend to agree that we do know the ordinary things that the sceptic tends to question. Conversational contextualism maintains that in order to question the truthfulness of an attribution sentence involving an ordinary proposition, the sceptic has to change the context and standards that are at play. However, how does the sceptic change the context, if he does? Proponents of conversational contextualism claim that he does so simply by introducing premise (1) of her argument (§1). Namely, when she asserts the first premise she raises the standards for attributing knowledge to the highest possible level, and does so by making the possibility that we are a BIV salient (a possibility that has to be eliminated in order for the subject to be considered as knowing a relevant proposition). The meaning of the word "know" shifts (Lewis et al., 1996). That leads us astray into believing that certain knowledge ascriptions conflict, when as a matter of fact they are compatible (Cohen, 2001: 89). Furthermore, we get a faulty impression that the principle of closure is not sustained. But, if the epistemic standards remain the same, according to

contextualists, then an attribution sentence involving an ordinary proposition and an attribution sentence including the denial of the sceptical hypothesis would have the same truth-values (see: Stine, 1976: 256; DeRose 1995: §9; Lewis, 1996: 564; Cohen, 1999: 67).

As we can see, the solution before us is quite refined and delivers on its promise to explain the intuitive force of each of the three propositions of the sceptical paradox. So far so good, but the question remains: how did we become so puzzled in the first place? In other words: if the sceptic is not actually denying what we have been claiming the whole time, why did we wrongly think that the sceptic raises a doubt about the truthfulness of our ordinary knowledge ascriptions?

According to conversational contextualists, this is because we are unaware of context-sensitive standards for the application of the word *know* and its cognates, and somehow imprudent as to the indexicality of these terms. This view represents the starting point of an *error theory* in epistemic contextualism that is commonly taken to imply that competent speakers are, to a certain extent, blind as to the semantic nature of the concept of knowledge (see: Hawthorne, 2004: 107; DeRose 2009: 159; cf. DeRose 1995: 40-1; Schiffer 1996: 325; Cohen 1999: 77; Stanley, 2005: 29). Error theory has been the target of many objections; such as Stephen Schiffer's argument that it is quite an implausible clarification of how we became "bamboozled by our own words" (Schiffer, 1996: 329); or MacFarlane's argument that it raises internal problems for the contextualist solution, namely, it casts some doubts on the plausibility of the contextualist's explanation of the everyday usage of the word "know," which in fact plays a crucial role in illustrating the thesis on the meaning of and on the manner in which we employ the verb "to know". (*see* MacFarlane, 2004; 2014: 181).[2] However, discussing all of these objections is beyond the scope of this paper. Therefore, we will focus on one particular kind of objection, namely Patrick Rysiew's (2001: 482-485) interpretation of Schiffer's critique.

According to Schiffer's vision of the contextualist application of error theory, contextualists cannot consistently claim that knowledge ascription sentences have semantics sensitive to conversational

---

[2] For other critics that aim to show that error theory is inappropriate to conversational contextualism, *see* Thomas Hofweber, 1999; Richard Feldman, 1999; John Hawthorne, 2004; Jason Stanley, 2004; John MacFarlane, 2005; Kent Bach 2005; Earl Conee 2005: 55 and others.

factors and maintain an error theory that he formulated in a way that it reduces to the following claim:

> /…/ people uttering certain knowledge sentences in certain contexts systematically confound the propositions their utterances express with the propositions they would express by uttering those sentences in certain other contexts. (Schiffer 1996:325)

Understood in this way, Schiffer argues that error theory rebuts the main contextualist thesis. Rysiew holds that in order to take a position on the merits of Schiffer's argument, we must first determine what it is exactly Schiffer has in mind with *error theory*. More precisely, Rysiew has noticed that we are in a position to find two types of error theory in Schiffer's formulation:

(i)   Error theory: competent speakers don't know what propositions are literally expressed by sentences containing context-sensitive terms.

(ii)   Error theory: competent speakers don't know what they mean when uttering certain context-sensitive sentences.

The first asserts that speakers are systematically puzzled as to what the proposition that their utterances literally express is. The second maintains that speakers are systematically confounded with regard to what they had in mind in uttering certain context-sensitive sentences. The majority of epistemologists seem to have understood this error theory under the first notion (*see*, for example, Cohen, 2005: 61; Feldman, 2001: 72-77; DeRose, 2009: 177). However, Rysiew has argued that contextualists are in fact forced to accept the second notion. What exactly does Rysiew have in mind?

As we may recall, "context", from the standpoint of conversational contextualists, is wholly grasped in terms of conversational factors (purposes, beliefs, interest, presuppositions, intentions, etc.). The meaning of the word "know" varies with changes of the context, i.e. with changes in these conversational factors. If that is the case, then the contextualist claim that we are systematically confused about the meaning of the word "know" becomes a claim that we are systematically confused about our own communicative intentions. Rysiew finds this extremely implausible. (Rysiew, 2001: 485). Indeed, even if contextualists find a way to explain how it is possible that we are not in a position to recognize our own communicative intentions, this

resolution would prove problematic. This is because (setting aside the fact that it stands in a deep discord with our linguistic competences) it seems that adopting the second version of the error theory would force us to re-visit the reasons for favouring the anti-sceptical strategy put forward by conversational contextualists in the first place. For we have taken, as one of the advantages of their strategy, that it is able to preserve our common-sense conviction that we do possess knowledge in many forms. Indeed, it would be odd to understand the goal of an anti-sceptical strategy in any other terms. This is what DeRose is getting at when he writes:

> /.../ our new solution is designed largely with the goal in mind of crediting most of our attributions of knowledge with truth. And no wonder. We in general take it as a strike against a theory of a common term of a natural language that it involves the speakers of that language in systematic and widespread falsehood in their use of that term. (DeRose, 1995, § 16)[3]

But is it not true that accommodating (ii) would burden the epistemic subject with too much error, which is something that we have strived to avoid from the very beginning? Rysiew points out that, if we were willing to maintain our everyday knowledge attribution at the cost of renouncing knowledge of our communicative intentions, in doing so we would end up swapping one brand of scepticism for another due to which we would "gain the world but lose our minds" (2001: 485). If Rysiew is right with respect to the consequences of the contextualists' adoption of the error theory, then it is natural to understand reliance on that theory as a crucial flaw of the conversational contextualist position.[4]

3   Inferential Thesis

As has already been mentioned, there are several differences between the two camps (conversational and inferential) on the path to reach the solution of the sceptical problem and, as we have said, we focus on only one aspect: the explanation of how we became puzzled with the sceptical argument in the first place.[5] In contrast to conversational contextualists, the explanation put forward by Michael Williams avoids reliance on an error theory altogether, instead putting

---

[3] See also: De Rose, 1999: 202; Cohen 1999: 83.
[4] A different opinion can be found in: Neta, R. 2003. "Contextualism and the Problem of the External World." *Philosophy and Phenomenological Research* 66 (1): 1–31
[5] See Duncan Pritchard, 2002, for a discussion on other *significant differences* between the compared theories.

forward a ''theoretical diagnosis'' of the sceptical argumentation and the closely related traditional epistemology (Williams, 2004a: 465). In order to understand why Williams opts for this approach, we must first examine how he understands the nature of context. For Williams, context is determined by the specifics of the subject matter we are dealing with and in light of which we judge whether one knows something. Within each thematic area, there exist certain presuppositions which are taken for granted and not questioned. Their role is to determine certain modes of justification, forms of explanations and conclusions in a particular context. Williams calls these kinds of standing assumptions *methodological necessities* (see: 2004a: 467; 2007:102).[6] If any one of these assumptions were to be questioned, the context of research would shift, and by virtue of which the context of knowledge ascription would shift accordingly. Here is Williams:

> Methodological necessities are standing presuppositions such that questioning them would lead one to question the competence of the form of inquiry they enable. In this way, they determine the disciplinary meta-context for such practices as historical. /…/ Raising general doubts about the usability of documentary and other historical evidence would not be not an especially rigorous approach to historical research, any more than entertaining skeptical doubts about the reality of the external world would be an exceptionally careful way of conducting experiments in physics. Rather, to bring up such issues changes the subject from history, or physics, to (a certain kind of) epistemology (which, as we are discovering, has disciplinary presuppositions of its own). (*Ibid*)

Accordingly, Williams argues that to change the context to epistemological (viz. to engage in sceptical epistemology) is not to raise the *standards* for *attributing knowledge, but* to change the subject of inquiry altogether. Worrying about the brain-in-a-vat possibility introduces a completely different kind of investigation: examination of sceptical epistemology, which is closely related to traditional epistemology. Traditional epistemology is unique in that it builds on the theoretical presupposition that beliefs can be grouped into epistemic kinds which stand in an invariant order of epistemic priority based on some broad characteristics of their content. According to this idea, certain beliefs (e.g. about perceptual appearances) establish an autonomous category of justified beliefs – usually called basic beliefs. The next (second) level is comprised of beliefs that are inferentially dependent on epistemically prior beliefs; at the third level there are beliefs that are

---

[6] According to that, Williams' account of contextual sensitivity of *knowledge-sentences* is to be found in variations (which are primarily the variations within the methodological assumptions and inferential forms) between the field of study that are the subject of knowledge in dissimilar contexts.

inferentially dependent on the previous one, and so on. This way of reasoning is built into the idea that our knowledge of the external world has to be based on some more primitive level of knowledge: experiential knowledge. According to Williams, the sceptic, just like traditional epistemologists, presupposes the priority of experiential knowledge over the knowledge of the external world. When she asserts the first premise, she already assumes that all our knowledge of the world depends on "experience'' as a generic source of knowledge. However, for Williams the idea that experiential knowledge is, in some general way, epistemologically prior to knowledge of the world, is an instance of "epistemological realism." Epistemological realists hold that our experiential knowledge of the outside world represents a natural epistemic kind which is, as such, in a certain objective relationship with other forms of knowledge. Amongst these relationships the most important one is the relationship of epistemic priority. However, according to Williams, epistemological realism is a theoretical assumption which, without a well-formulated reason, we are not obliged to accept.[7] Furthermore, Williams argues that the epistemic requirement for knowledge possession itself is context-dependent, an idea quite at odds with the notion that there exists a general objective epistemic criterion for knowledge or ultimately standards of justification. On Williams' account, knowledge and justification are in some ways relative to context and hence, not "invariant" or "absolute." Understandably, the same rationale can be applied to the idea of epistemic standards. Different contexts merely impose different epistemic standards and in no way is the epistemic context hierarchically superior to our everyday epistemic context.[8] Thus, Williams frequently insists on the following:

> The skeptic takes himself to have discovered, under the conditions of philosophical reflection, that knowledge of the world is impossible. But in fact, the most he has discovered is that knowledge of the world is impossible under the conditions of philosophical reflection. (Williams, 1991: 130; cf. 2001: 3-4; 2004b: 316).

---

[7]This way of partitioning our beliefs into privileged and problematic classes is the central idea of *foundational theories* of knowledge in traditional epistemology, the primary aim of which is to provide answers to further questions, such as how our beliefs of the external world are justified by experiential belief (or in other words) to "recognize a range of distribution principles, which transmit positive epistemic status from basic to non-basic beliefs." (Wiliiams 2007: 96)

[8] Whether this raises objectionably relativistic outcomes is a matter of a wide-ranging debate. See Sosa, E. and J. Kim (eds.): 2000, Williams: 2007; *Rysiew: 2011; and others.*

This is the general contextualist manoeuvre according to which the impossibility of knowledge in an epistemological context does not entail the impossibility of knowledge in an everyday, ordinary context. Thus, laymen knowledge is isolated from sceptical undermining and the basic contextualistic strategy is ensured. Moreover, if Williams' diagnosis of the source of the sceptical problem is correct, then philosophical scepticism carries on its back a serious burden – to ensure positive authorization of epistemological realism.[9] Furthermore, as a result of this insight, we can say that the sceptic's doubts are not *intuitive* or *natural* as they seemed at first sight, rather they depend on theoretical ideas about knowledge and its cognates, something that we are not obligated to accept.[10] Until we are shown differently, it is far from obvious that these epistemological presuppositions force us into thinking about knowledge or justification in this particular way. And this manoeuvre of shifting of the burden of theory is, dialectically taken, a noteworthy feature of *any anti-sceptical* strategy whatsoever.

4   Concluding Remarks

We have provided a brief overview of the philosophical problem of scepticism and touched upon the two main contextualist anti-sceptical strategies. While the conversational contextualist approach ultimately relies on a highly problematic error theory which entails that we are all constantly mistaken about our communicative intentions, we have shown that, at least in this regard, Williams' inferential position clearly has a better starting position to provide an effective contextualist solution to the sceptical problem. In addition, we have seen that the inferential approach centres on shifting the burden of the theory, which further contributes to its potential as an effective anti-sceptical solution.

---

[9] This idea stands in sharp contradiction with another presupposition of traditional epistemology, the so called *claimant-challenger asymmetry* (*see* Williams 2007), according to which the mere possibility of error lays the justificatory burden on the claimant. However, because challenges are not always legitimate (rather challengers subject to contextual factors) according to Williams, this traditional assumption should be rejected.

[10] Why the question of *intutivness* and *naturalness* of sceptical doubts matters see: Williams, 1991, 2004a *passim*.

References

Bach, K. (2005). "The Emperor's New 'Knows'." In G. Preyer & G. Peter (Eds.), *Contextualism in Philosophy: Knowledge, Meaning, and Truth.* Oxford University Press, pp. 51–8.

Blome-Tillman, M. (2008). "The Indexicality of 'Knowledge'." *Philosophical Studies*, *138* (1), pp. 29–53.

Cohen, S. (1988). "How to Be a Fallibilist." *Philosophical Perspectives* 2, pp. 91–123.

Cohen, S. (1999). "Contextualism, Skepticism, and the Structure of Reasons." *Philosophical Perspectives*, 13, pp. 57–90.

Cohen, S. (2005) "Contextualism Defended." In Steup, M. and Sosa, E. (Eds.), *Contemporary Debates in Epistemology*, Oxford: Blackwell, pp. 56–62.

Conee, Earl. (2005). "Contextualism Contested." In Steup, M. and Sosa, E. (Eds.), *Contemporary Debates in Epistemology*, Oxford: Blackwell), pp. 47-56.

DeRose, K. (1995). "Solving the Skeptical Problem." *Philosophical Review*, 104, pp. 1–52.

Feldman, R. (1999). "Contextualism and Skepticism." *Philosophical Perspectives*, 13 pp. 91–114.

Heller Mark. (1999). "Relevant Alternatives and Closure." *Australasian Journal* of *Philosophy*, 77 (2), pp. 196–208.

Hofweber, T. (1999). "Contextualism and the Meaning-Intention Problem." In K. Korta, E. S. and Arrazola, J. (Eds.), *Cognition, Agency and Rationality*, pp. 93–104.

Howthorne, J. (2004). *Knowledge and Lotteries*. Oxford: Calendron Press.

Kennedy, C. (1999). *Projecting the adjective: The syntax and semantics of gradability and comparison*. NewYork: Garland. [Santa Cruz: University of California, Santa Cruz dissertation, 1997.

Lewis, D. (1996). "Elusive Knowledge." *Australasian Journal of Philosophy*, 74 (4), pp. 549–567.

MacFarlane, J. (2005). "The assessment sensitivity of knowledge attributions." *Oxford Studies in Epistemology*, 1, pp. 197–233.

MacFarlane, J. (2014). *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford: Oxford University Press.

Neta, R. (2003). "Contextualism and the Problem of the External World." *Philosophy and Phenomenological Research* 66, (1), pp. 1–31.

Pritchard, D. (2002). "Two Forms of Epistemological Contextualism." *Grazer Philosophische Studien 64*, pp. 19–55.

Rysiew, P. (2001). "The Context-Sensitivity of Knowledge Attributions." *NOÛS* 35 (4), pp. 477–514.

Rysiew, P. (2011). "Relativism and Contextualism." In Hales, S. D. (Ed.), *A Companion to Relativism*, Blackwell, pp. 286-305.

Schiffer, S. (1996). "Contextualist Solutions to Skepticism". *Proceedings of the Aristotelian Society*, 96, pp. 317–33.

Sosa, E. and J. Kim (Eds.). (2000). *Epistemology: An Anthology*. Blackwell, Oxford.

Stanley, J. (2005). *Knowledge and Practical Interests*. Oxford University Press.

Stine, G. C. (1976). "Skepticism, Relevant Alternatives, and Deductive Closure." *Philosophical Studies*, 29 (4), pp. 249–261.

Unger, P. (1975). *Ignorance: A Case for Skepticism*. Oxford: Oxford Univeristy Press.

Williams, M. (1991). *Unnatural Doubts: Epistemological Realism and the Basis of Skepticism.* Basil Blackwell, Oxford.

Williams, M. (2001). "Contextualism, Externalism and Epistemic Standards." *Philosophical Studies*, 103, pp. 1–23.

Williams, M. (2004a). "Skepticism and the Context of Philosophy." *Philosophical Issues*, 14, pp. 456–75.

Williams, M. (2004b). "Knowledge, Reflection and Skeptical Hypotheses." *Erkenntnis*, 61, pp. 315–343.

Williams, M. (2007). "Why (Wittgensteinian) Contextualism Is Not Relativism." *Episteme*, 4, pp. 93-114.

Lalit Saraswat

Indian Institute of Technology Bombay

Bridging Selection Theory and Autopoietic Epistemology:

Explicating 'Downward Causation'

Abstract

Selection theory of knowledge finds its support in Popper and Campbell's conception of Blind-variation-and-selective-retention BVSR, trial-and-error-elimination (TEE) method, and nested hierarchy model. It is argued for the centrality and sufficiency of 'selection theory' in explaining knowledge in all life forms and their activities. It has its main feature in addressing the notion of emergence and the transition of thought processes at nested hierarchical levels. BVSR, along with TEE, claims to have an explanatory power to address the phenomenon at each level of life – organic, individual, socio-cultural. In the hierarchically functional 'downward causation', the laws of higher levels govern the events in the lower levels and autopoiesis is envisaged. TEE holds good for the fallibilistic epistemology incorporating 'learning from our own errors', and also accounts for the growth of knowledge. Popper (1972) incorporates TEE in a tetradic schema (P_1 » TT » EE » P_2) [where P is a problem, TT is a tentative theory, EE is error elimination]. Each level of hierarchy (nested) involves such schematic problem-solving. BVSR and TEE, under selection theory, offer creative solutions by continuously finding problems, generating a number of *blind* variants, and minimizing the search spaces by selection.

Keywords: BVSR, TEE, autopoiesis, selection theory of knowledge, hierarchical organization, emergence

0   Introduction

Selection theory identifies natural selection as one prime case of it; another is from the developmental and autopoietic features that are claimed to have no apparent connection with natural selection. Selection theory, Blind-variation-and-selective-retention (BVSR), Trial and error elimination (TEE), and 'Downward causation' play an inevitable role in hierarchically organised systems where feedback recursive strategies are performed with internal and external selection criteria. Downward causation integrates both the selection theory approach and the autopoietic features of self-organising system. Such causal systemic understanding and features act as a bridge to fill the gap between pure selectionism and autopoietic characteristics. This is all done under natural selection framework and leads to epistemic and cognitive growth. Such mechanisms help in multi-level variations and selections while retaining the progressively intensified form of stability and autopoietic functioning. The evolution by natural selection is mostly seen as reductionistic in nature, but some evolutionary epistemologists hold a reductionist viewpoint even if they are not 'microparticulate-derivationists'. Instead, many of them qualify as emergentists (Campbell, 1974: 189). The mistakes done by reductionistic approaches, mostly by determining the denial of a few significant points made by vitalists, should be avoided. Denial of design is not what Darwin must have sought; it is perhaps 'the argument from design' where some intelligent designer is posited; this could also be observed in the case of acceptance of 'teleonomy' without teleological explanatory understanding. The below mentioned two basic tenets allow reductionists to adhere with hierarchically organized biological systems. These are then merged with the two principles of emergentism and 'downward causation'. With this addition, a complete picture could be grasped for the overall understanding of emergence, regulation, and self-organizing aspects in hierarchical systems. Otherwise, without such a comprehensive understanding, the reductionists and the vitalists seem to have an incomplete or incoherent picture of reality. Campbell's contribution is noteworthy in bringing a meeting point for these two schools of thought.

The reductionistic points are:

> i) All processes at the higher levels are restrained by an act in conformity to the laws of lower levels, including the levels of subatomic physics.
>
> ii) The teleonomic achievements at higher levels require for their implementation specific lower-level mechanisms and processes. Explanation is not complete until these micro-mechanisms have been specified. (Campbell, 1974: 180)

However, Campbell outlines another two modified points to accommodate the emergentism, complexities, and reproductively available collective units. He notes two further points as:

> iii) (The emergentist principle) Biological evolution in its meandering exploration of segments of the universe encounters laws, operating as selective systems, which are not described by the laws of physics and inorganic chemistry, and which will not be described by the future substitutes for the present approximations of physics and inorganic chemistry.

> iv) (Downward causation) Where natural selection operates through life and death at a higher level of organisation, the laws of the higher-level selective system determine in part the distribution of lower-level events and substances. Description of an intermediate-level phenomenon is not completed by describing its possibility and implementation in lower-level terms. Its presence, prevalence or distribution (all needed for a complete explanation of biological phenomena) will often require reference to laws at a higher level of organisation as well. Paraphrasing Point 1, all processes at the lower levels of a hierarchy are restrained by and act in conformity to the laws of the higher levels. (Campbell, 1974: 180)

The converging aspects of reductionistic and vitalistic viewpoints provide interesting features by addressing the issues related to the overall understanding of lower level physico-chemical, law-bound conformities (micro-mechanistic), and higher-level laws governing the lower-levels with continuous referential and feedback nodal points.

It can be explained with the help of an example of jaws of termite, pointing out the conformational points between the reductionist-antireductionist viewpoints on this. It is about the intricate

conformity of laws of physics, DNA-protein combinatorial mechanisms performing a task under selection pressures. There have been instances from immune system and tumor controlling immunological coding-decoding processes, which also represent the 'downward causation.' The indications of such causation, boundary conditions,[1] and dual control, are observed in Polanyi's (1968) seminal paper "Life's Irreducible Structures." Evolutionary principles progressively intensify the higher principles that in turn harness the lower levels (even to physico-chemical topographical arrangements). We notice that selective pressures play important roles at each stage and each boundary where each semi-stable template gets orders for the subsequent level to be addressed further. Epistemological relevance through natural selection provides the hypothetical, modest attitude towards knowledge. With a thoroughgoing neo-Darwinian evolutionary epistemological approach, the epistemic humility is envisaged, while recognising an indirect presumptive perception and science. "We do not know reality in the *Ding an Sich's* own language." (Campbell, 1991: 184) There is a need to rephrase the metaphors we have been using to describe reality. There is not a literal language for such descriptions of truth. Even mathematics, as a language, is a metaphorical take on reality, not a literal one. This lesson is one of the key points of learning from evolutionary considerations.

Henri Poincare is recognized as the proponent of the selectionistic theory, as Campbell sees the Poincarean notion of 'mathematical beauty' as a vicarious selector. Campbell defies Barham's (1990) position of considering Poincare as an anti-selectionist. Campbell outlines an interesting point in both evolutionary epistemology and biology about the unit of selection and criteria to determine it. BVSR acts as a bridge to fill the gap between pure selectionism and autopoietic features. It features as a strong point of collaboration between autopoiesis and 'selection-by environment'. Though Barhamian perspective is quite a significant one for the age-old issue of 'unit of selection'. Barham's points on autopoietic understanding along with Campbellian 'downward causation' play an important role in addressing the issue of complexity, order, and different units of selection. 'Downward causation' has been championed and validated by Campbell (1974) and Polanyi (1968). Its central tenet is that higher stages have strategic control

---

[1] There are two types of boundary conditions governing at hierarchical organisations; one is machine type – which is design oriented almost downwardly caused functionality, and the second one is a test-tube type where boundaries are incorporated by the experimenter.

over lower ones even in the complex hierarchized form of the system. The notion of such causation has its implication on an understanding of emergence. An internal selection criterion under selection theory is developed for the understanding of evolution of cognition and knowledge. The concept of internal selection is endorsed for the fitment. The fitment between an organism and environment has an intricate aspect of an internal selection mechanism that is handled by autopoietic needs of the system. It propagates and limits the overall dynamics along with other selection pressures. It is suggested that the criteria of selectors – internal or external for the fitment, self-organising stabilities with feedback recursive strategies and even teleonomic explanations (apparent purpose, goal) – is easily achieved by non-clairvoyant blind variations and selective retentions. As it may be noticed with naturalistic evolutionary framework, punctiform certainty of belief or knowledge is never claimed, but is rather decreed not to adhere to such traditional foundationalist, infallible epistemic structures.

The demarcation between a living and non-living had been a central issue for biologists and philosophers. The unique feature of the living system is noticed and discussed in detail, a phenomenon that is the self-organisational property of a system called 'Autopoiesis'. This, according to Varela and Maturana (1980), is the unique classifying feature of living systems. It could be said that evolutionary constraints play quite a significant role in shaping the epistemic and cognitive aspects of organisms. These constraints are noteworthy, and an attempt has been made to point out specific issues of possibility of (i) inquiry, (ii) discovery, (iii) emergence in hierarchized systems via control theory that is *plastic* in nature, (iv) notion of 'downward causation' that becomes a key feature in identifying the self-organising property and converging the reductionistic and anti-reductionistic approaches, (v) the unique feature of 'autonomic' structures, (vi) a significant aspect of 'autopoiesis' in living systems, and (vii) behavioural novelty, creativity. Popper and Campbell both majorly complement each other in developing the 'epistemology of other' that is the 'selection theory of knowledge'. Their method of TEE and BVSR are closely connected to address and explain the origin and growth of knowledge. Campbell's nested hierarchical ten stages under BVSR attain the kind of autonomical and creative status.

1   The autopoietic evolutionary theory of knowledge

A hierarchized self-organised living system could be attributed with feature that "the autopoietic requirement is presented as producing limitations, and systematic imperfections, on adaptation or knowledge." (Campbell, 1991: 171) Such self-organizational autopoietic epistemic and cognitive structure *functions* as an autonomic body with hierarchically organised systems and subsystems with a *feedback*, and perhaps feedforward controls. These epistemic systems work under the principle of 'downward causation' (higher level influencing lower levels in feedback manner) and give rise to an interesting point for *emergence* of epistemic, cognitive, and regulatory controls. Living systems, in this scenario, are considered as open, hierarchically organized systems consisting of interrelated linked parts with feedback regulatory principles. A few key features about the multi-level causation model with organismal constraints and form are also outlined. Living systems are multi-level systems where parts and whole both determine and constrain the structure and function for a flux of cause in effect in two directions – upward and downward levels of organizational hierarchy. And a way to define an autopoietic living system is:

> Autopoietic system is organized as a bounded network of processes of production, transformation and destruction of components which:
>
> (i) Through their interactions and transformations continuously regenerate and realize the network of processes that produced them
>
> (ii) Constitute the system as a concrete entity in the space in which the components exist by specifying the topological realization of the system as such a network. (Varela and Maturana, 1980)

It is emphasized that the notion of constituents, orderliness, and synchronization amongst subsystems in a living system are crucial features for knowledge and cognition. The notion of telos, purpose, attached to any autopoietic living or physical systems is undermined. Maturana asserts, "If a composite unity is structurally plastic, then adaption as a process of structural coupling to the medium that selects its path of structural change is a necessary outcome." (Maturana, 1981: 29) These systems are *open* systems and get well adapted with the environmental selection pressures and they couple themselves accordingly. Such coupling of

organisms/systems with their environment is in part deterministic and in part indeterministic, in other words, they are *plastic* in nature (plasticity is discussed in a later section).

Popper (1978), in a similar line of thought, addresses the emergence of mind and 'downward causation' while depicting natural selection as the primary force. One of the important points about this 'roundabout method of selection' is that it throws light on the problem of downward causation, to which Donald Campbell has extensively contributed. Along with causation that is downward in nature and hierarchically controlled, Popper emphasizes the chance like situations and indeterminacy in 'downward causation'.

> I think that these considerations tell us a lot about natural selection. While Darwin still worried that he could not explain variation, and while he felt uneasy about being forced to look at it as chance like, we can now see that the chance like character of mutations, which may go back to quantum indeterminacy, explains how the abstract invariances of the environment, the somewhat abstract selection pressures, can, by selection, have a downward effect on the concrete living organism — an effect that may be amplified by a long sequence of generations linked by heredity. (Popper, 1978: 348)

Popper calls attention to the difficulty in grasping how downward causation works, as most of the time 'upward causation' seems to capture scientific minds easily. Downward causation shows how higher structures formed could govern or influence the lower level structures as well as substructures. This all is done under selective pressures.

> The difficulty of understanding downward causation is this. We think we can understand how the substructures of a system cooperate to affect the whole system; that is to say, we think that we understand upward causation. But the opposite is very difficult to envisage. For the set of substructures, it seems, interacts causally in any case, and there is no room, no opening, for an action from above to interfere. It is this that leads to the heuristic demand that we explain everything in terms of molecular or other elementary particles. (Popper, 1978: 48)

Popper tells us that 'upward causation' usually employs the notion of 'reductionism'. He gives credit to Darwin for the theory of natural selection that allows handling issues related to complex

phenomena like free will, the emergence of mind, and diversity. His own understanding and position from earlier standpoints on the growth of knowledge and evolution is transformed after noticing the significance of natural selection principles and 'downward causation.'

> [D]ownward causation can sometimes at least be explained as selection operating on the randomly fluctuating elementary particles. The randomness of the movements of the elementary particles — often called "molecular chaos" — provides, as it were, the opening for the higher-level structure to interfere. A random movement is accepted when it fits into the higher level structure; otherwise it is rejected. (Popper, 1978: 48)

Therefore, it can be understood that the indeterminate structuralizations and physicochemical conformities are selected by the higher level and ordered in an organized form under selective feedback strategies.

2   A hierarchical system of the plastic control

The aim of the nested hierarchy is to identify the selective and retentive systems at each evolutionary level. It is core to the evolutionary theory of knowledge under blind variation and selective retention model. The notion of 'nested' implies the vicariously available explorations, which are provided by presumed mechanisms. Evolutionary mechanisms applied to primitive and complex life forms involve various levels and stages ontogenetically and phylogenetically. Such hierarchical analogy could also be envisaged for the growth of human knowledge. These hierarchies are in a continuum and employ selectively retentional processes. Popper comments on the possible hierarchical systems:

> /…/ it consists of a certain *views of evolution* as a growing hierarchical system of plastic controls, and of at certain view of organisms as -incorporating- or in the case of man, evolving exosomatically— this growing hierarchical system of plastic controls. The Neo-Darwinist theory of evolution is assumed; but it is restated by pointing out that its "mutations" may be interpreted as more or less accidental trial-and-error gambits, and "natural selection" as one way of controlling them by error-elimination. (Popper, 1966: 23)

Popper's (1972) sketch of physical, biological, and social systems suggests quite a remarkable picture of their functioning and structuralization. The indeterministic features of physical systems may be classified as cloudy and possessing irregularity, disorder, and, unpredictability. And the deterministic physical system as a clock that is reliable, precise, full of regularities, orderliness, and of a high predictive value. Various natural processes and phenomena could be examples of such a classification, depending upon the degree of randomness or order. According to Popper, plants and animals would be put more towards the clock and cloud side respectively. Furthermore, infants would be placed more towards the cloud than an adult. He uses an example of two-winged flies and its cluster, flying as a 'whole' and acting more like an irregular cloud, but with some orderliness simultaneously, as could also be observed in the case of disorderly molecular motions in a gaseous cloud or a storm. Popper (1972) takes a position against the notions of Newtonian (though he may be classified as a dissenter of rgw deterministic approach on his notion about solar system) and deterministic viewpoints. It is held that 'all clouds are clocks-like; even the most cloudy of clouds' (Popper, 1972: 212). For a physical determinist there would be no possibility of having a *cloudy* system, as everything has a clock as its deterministic foundation. Therefore, out of our limited sense of nature's functioning, clouds are placed apart from clocks, while even clouds have a clock like arrangement if given proper notice.

It is significant to account for how a deterministic perspective may prevail, yet it doesn't mean that indeterminism is bad or allows a system to be non-functional. Popper mentions that the Peircean viewpoint as follows: "/.../ at any rate we could not possibly claim to know, from experience, of anything like a perfect clock, or of anything even faintly approaching that absolute perfection which physical determinism assumed." (Popper 1972: 212) This conjecture introduces a notion and role of *chance*, as against the notion of a deterministic perspective. It makes an interesting hybrid of clocky-clouds or cloudy-clocks, introducing the scope of randomness along with the laws of chance and statistical probabilities. The biological perspective, scope, and consideration associated with the newly formed physical indeterminacy is further explored. There is a difference of openness and closedness while describing a system deterministically or indeterministically. Biological systems could thus be termed as more open and indeterministic in nature. But it can also be said that the closed system could be observed to behave as indeterministic. Popper distinguishes between   philosophical/psychological determinism and

physical determinism; stating that the former one is quite vague in its formulations when it says thing like 'Like effects have like causes', or even 'Every event has a cause'. Furthermore, this notion of relation between effect and cause doesn't say much about the *precision* in question. This is why the imperfect and incompatible relation between observables and measurables leave room for indeterminism.

Popper, being an indeterminist, warns that the deterministic position hampers the idea of possibilities of creative knowledge, creative processes, and novelty. It is so because the Laplacean deterministic demon (reality is physically and completely deterministic in nature to minutest details) would never allow having any non-physical characteristics like *aims*, *expectations*, *dispositions*, and *purposes*. For him, indeterminism prevails – this can also be observed if evolution is true. It would be of great interest to know how the physical system or the world would have arisen or produced the non-physical features like grammar, rules, abstract entities, reasons. Also, how do these evolutionary indeterministic products start to influence the physical universe? This is the classic case of the Popperian three worlds (physical, mental, and world of objective knowledge) interaction, where the emergent evolutionary products attain at least partial autonomy. The formulation of the three worlds hypothesis connects Popper's proposal to Plato's metaphysical epistemology. However, Popper's 'third world' is an autonomous (partially), man-made, evolutionary, and revisionistic in nature, as opposed to Plato's. There is a unique interactive relationship between the proposed three worlds. 'World 1' (physical) and 'World 3' (theories, arguments, books etc.) are connected via an intermediary 'World 2' (mental). The products or structures are always more important than the problem that produces those products. "[W]e should realise that the category of problems, those concerned with the products in themselves, is in almost every respect more important that the first category, the problems of production." (Popper 1972: 194) It could be noted that the evolutionary products have tendency to transcend themselves from their constituting elements; and that the higher level might in turn influence them. With the third world and its products 'we persistently act, and have influence upon it and are influenced by it.' The third world is *autonomous*.

Such indeterminism is quite important for problem-solving. But a puzzling point arises when Popper puts Schlick's (1925), a determinist, idea that raises a point that chance-like situations could be dangerous if indeterminism is held. "[F]reedom of action, responsibility, and mental

sanity, cannot reach beyond the realm of causality: they stop where chance begins /…/ a higher degree of randomness /…/ [simply means] a higher degree of irresponsibility." (Popper 1972: 226) It could be observed that how things progress from completely closed, deterministic systems to chance-like but intermediary systems with freedom and control, could be envisaged together. Such is the notion of 'plastic control' as opposed to the 'cast-iron' one. Also, we can allude to Popper's position that for resolving the issues of *determinism* & *indeterminism* together, a 'master-switch' like phenomenon would also be unsatisfactory, as it is again a form of a snap-decision. But the downward causation viewpoint with selectional criteria at various hierarchical levels provides a hope in deciphering the status of such systems, where intermediary stages have a significant role at upper and lower levels. Evolutionary emergent features get explanatory as well as theoretical power from Popper's schema to address issues like novelty, emergence of newer problems, etc. Therefore, a comprehensive picture of the growth of knowledge is achieved only under evolutionary principles.

3   Concluding remarks

'Selection theory' has the primacy, centrality, and sufficiency in explaining knowledge in all life forms and their activities. It has its main advantage in addressing the notion of emergence and transition of thought processes at nested hierarchical levels. Our aim is to grasp the ideas mainly propounded by Donald Campbell and Karl Popper, envisaging the broadest, most epistemological pursuit. The selection theory is espoused to encompass the possibility of adaptive learning, knowledge acquisition processes, scientific theory making, and creativity. The most crucial feature of selection theory of knowledge is to grasp the concept of *blind* variations and selective retention (BVSR) and the notion of downward causation. BVSR, along with the trial and error elimination (TEE), claims to have an explanatory power to address the phenomena at each level of life (organic, individual, socio-cultural). Selection theory of knowledge is merged with the hierarchically functional 'downward causation', where the laws of higher levels in the selective system govern the events in the lower levels. TEE holds good for the fallibilistic epistemology, incorporating the idea of 'learning from our own errors', and accounts for the growth of knowledge. BVSR and TEE under selection theory offer numerous variations, selective retention, and creative solutions to the already existing problems and explore next level problems and so on.

In all cases where the genuine increase of knowledge is envisaged, BVSR and TEE play a crucial role in some or another way; e.g. generating the huge number of *blind* variants, minimizing the search spaces with selection (may be blindly in some cases), retain, and reproduce. Apart from naturally occurring phenomena under BVSR and TEE, the application of this model ranges from evolutionary computing, genetic algorithm, artificial intelligence, decision-making, creativity studies, etc.

References

Barham, J. (1990). "A Poincarean approach to evolutionary epistemology". *Journal of Social and Biological Structures*, *13* (3), pp. 193–258.

Campbell, D. T. (1974). "Downward causation'in hierarchically organised biological systems". In Ayala, F. J. and Dobzhansky, T. (Eds.), *Studies in the Philosophy of Biology*, pp. 179 – 186.

Campbell, D. T. (1988). "The Author Responds: Popper and Selection Theory". *Social Epistemology*, 2 (4), pp. 371–377.

Campbell, D. T. (1991). "Autopoietic evolutionary epistemology and internal selection". *Journal of Social and Biological Structures*, *14* (2), pp. 166–173.

Maturana, H. (1981). "Autopoiesis//Autopoiesis: A theory of living organization". *Psychology and Biology*, (2), pp. 21–33.

Maturana, H., & Varela, F. (1980). "Autopoiesis and Cognition". *Boston Studies in the Philosophy and History of Science*, *42*.

Polanyi, M. (1968). "Life's Irreducible Structure: Live mechanisms and information in DNA are boundary conditions with a sequence of boundaries above them". *Science*, *160* (3834), pp. 1308–1312.

Popper, K. (1963). *Conjectures and Refutations*. London: Routledge and Kegan Paul.

Popper, K. (1978). "Natural selection and the emergence of mind". *Dialectica*, *32* (3-4), pp. 339–355.

Popper, K. R. (1966). "Of Clouds and Clocks an Approach to the Problem of Rationality and the Freedom of Man". In *Objective Knowledge*. St. Louis, Missouri: Washington University.

Popper, K.R. (1972). *Objective knowledge: An Evolutionary Approach.* Oxford: Clarendon Press.

Popper, K.R. (1979). *Objective knowledge: An Evolutionary Approach* (Revised Edition). Oxford: Clarendon Press.

Quine, W. V. O. (1969). *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Schlick, M. (1925). *Naturphilosophie*. Ullstein.

Wuketits, F. M. (2006). "Evolutionary epistemology: The Non-Adaptationist Approach". In Gontier, N., Bendegem, J. P., Aerts, D. (Eds.), *Evolutionary Epistemology, Language and Culture*. Springer, pp 33–46.